RATE OF CONVERGENCE OF WONG-ZAKAI APPROXIMATIONS FOR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

ISTVÁN GYÖNGY AND PABLO RAÚL STINGA

Abstract. In this paper we show that the rate of convergence of Wong-Zakai approximations for stochastic partial differential equations driven by Wiener processes is essentially the same as the rate of convergence of the driving processes W_n approximating the Wiener process, provided the area processes of W_n also converge to those of W with that rate. We consider non-degenerate and also degenerate stochastic PDEs with time dependent coefficients.

1. Introduction

Consider for each integer $n \geq 1$ the stochastic PDE

$$du_n(t,x) = (L_n u_n(t,x) + f_n)\,dt + (M_n^k u_n(t,x) + g_n^k)\,dW_n^k(t), \tag{1.1}$$

for $(t,x) \in H_T = (0,T] \times \mathbb{R}^d$, for a fixed $T > 0$, with initial condition

$$u_n(0,x) = u_{n0}(x), \quad x \in \mathbb{R}^d, \tag{1.2}$$

given on a probability space $(\Omega, \mathcal{F}, P)$, where L_n and M_n^k are second and first order differential operators in $x \in \mathbb{R}^d$, respectively for every $\omega \in \Omega$. The free terms, f_n and $g_n = (g_n^k)$ are random fields, and $W_n = (W_n^k)$ is a continuous d_1-dimensional stochastic process with finite variation over $[0,T]$, for $k = 1, \ldots, d_1$.

Unless otherwise stated we use the summation convention with respect to repeated indices throughout the paper. The summation convention is not used if the repeated index is the subscript n.

The operators L_n, M_n^k are of the form

$$L_n = a_n^{ij}(t,x) D_i D_j + a_n^i(t,x) D_i + a_n(t,x), \qquad M_n^k = b_n^{ik}(t,x) D_i + b_n^k(t,x),$$

where $a_n^{ij}, \ldots, b_n^k$ are real-valued bounded functions on $\Omega \times [0,T] \times \mathbb{R}^d$ for all $i,j = 1,...,d$, $k = 1,...,d_1$, and integers $n \geq 1$, $D_i = \frac{\partial}{\partial x^i}$ for $i = 1,2,...,d$, and x^i is the i-th co-ordinate of $x \in \mathbb{R}^d$. The free terms, f_n, g_n^1,...,$g_n^{d_1}$ are real-valued functions on $\Omega \times [0,T] \times \mathbb{R}^d$ for each n. We assume that L_n is either uniformly elliptic or degenerate elliptic for all n.

Assume that the operators L_n, M_n^k, the free terms f_n, g_n^k and the initial data u_{n0} converge to some operators

$$L = a^{ij}(t,x) D_{ij} + a^i(t,x) D_i + a(t,x), \quad M^k = b^{ik}(t,x) D_i + b^k(t,x), \tag{1.3}$$

random fields f, g^k and initial data u_0 respectively, and $W_n(t)$ converges to a d_1-dimensional Wiener process in probability, uniformly in $t \in [0,T]$. Then under some smoothness conditions on the coefficients of L_n, L, M_n^k, M^k and on the data u_{0n}, f_n, g_n^k, u_0, f, g^k, and under some additional conditions on the convergence of the related area processes and on the growth of the auxiliary process B_n (defined in (2.2) and (2.3) below), the solution u_n to (1.1) converges in probability to a random field u that satisfies the stochastic PDE

$$du(t,x) = (Lu(t,x) + f)\,dt + (M^k u(t,x) + g^k) \circ dW^k(t), \quad (t,x) \in H_T \tag{1.4}$$

2010 *Mathematics Subject Classification.* 60H15, 60H35, 93E11.

Key words and phrases. Wong-Zakai approximation, nonlinear filtering.

The second author was supported by Ministerio de Ciencia e Innovación de España MTM2008-06621-C02-01 and Grant COLABORA 2010/01 from Planes Riojanos de I+D+I.

with initial condition

$$u(0,x) = u_0(x), \quad x \in \mathbb{R}^d, \tag{1.5}$$

where '$\circ$' indicates the Stratonovich differential. (See, e.g., [8] and [9].) When M^k and g^k do not depend on the variable t, then

$$(M^k u(t,x) + g^k) \circ dW^k = \tfrac{1}{2}(M^k M^k u(t,x) + M^k g^k)\,dt + (M^k u(t,x) + g^k)\,dW^k.$$

One of the important questions in the analysis of approximation schemes is the estimation of the speed of convergence. In this paper we show that, if the continuous finite variation processes W_n and their area processes converge almost surely to a Wiener process W and to its area processes, respectively, with a given rate, then $u_n(t)$ converges almost surely with *essentially* the same rate. The results of this paper are motivated by a question about robustness of nonlinear filters for partially observed processes, $(X(t), Y(t))_{t\in[0,T]}$. For a large class of *signal and observation* models, the signal X and the observation Y are governed by stochastic differential equations with respect to Wiener processes, and a basic assumption is that the signal process is a non-degenerate Itô process. Thus the signal is modelled by a process, which has infinite (first) variation on any (small) finite interval. In practice, however, due to the smoothing effect of measurements, the "signal data" is a process which has finite variation on any finite interval. This process can be viewed as an approximation Y_n to Y, and it is natural to assume that Y_n and its area processes converge almost surely in the sup norm to Y and its area processes, with some speed. By a direct application of the main theorems of the present article one can show that the "robust filtering equation", with Y_n in place of Y, admits a unique solution p_n which converges almost surely with almost the same order to the conditional density of $X(t)$ given the observation $\{Y(s) : s \in [0,t]\}$. The filtering equations in case of correlated signal and observation noise are stochastic PDEs with coefficients depending on the observations. Thus approximating the observations we approximate also the differential operators in the stochastic PDEs. This is why we consider equation (1.1) with random operators L_n and M_n^k depending also on n, the parameter of the approximation.

Our results improve and generalise the results of [12] and [17], where only half of the order of convergence of W_n is obtained for the order of convergence of u_n. Moreover, our conditions are weaker, and we prove the optimal rate also in the case of degenerate stochastic PDEs, which allows to get our rate of convergence result also in the case of degenerate signal and observation models.

Wong-Zakai approximations of stochastic PDEs were studied intensively in the literature. See, for example, [1]-[9], [12]-[14], [17]-[18], and the references therein. With the exception of [4], [14], [12] and [17] the papers above prove convergence results of Wong-Zakai approximations for stochastic PDEs with various generalities, but do not present rate of convergence estimates. Wong-Zakai type approximation results for semilinear and fully nonlinear SPDEs are obtained via rough path approach in [5]–[7].

In [4], the initial value problem (1.1)–(1.2) is considered with non-random coefficients and without free terms, when W_n are polygonal approximations to the Wiener process W. By the method of characteristics it is proved that $u_n(t,x)$ converges almost surely, uniformly in $(t,x) \in [0,T] \times \mathbb{R}^d$. Though the rate of convergence of u_n to u is not stated explicitly in [4], from the rate of convergence result proved in [4] for the characteristics, one can easily deduce that for every $\kappa < 1/4$ there exists a finite random variable ξ_κ such that almost surely $|u_n(t,x) - u(t,x)| \le \xi n^{-\kappa}$ for all $t \in [0,T]$ and $x \in \mathbb{R}^d$. We note that for polygonal approximations the almost sure order of convergence of W_n and its area processes are of order $\kappa < 1/2$, and thus by our paper the almost sure rate of convergence of the Wong-Zakai approximations is the same $\kappa < 1/2$, in Sobolev norms, and via Sobolev's embedding in the supremum norm as well. In [14] the rate of convergence of Wong-Zakai approximations of stochastic PDEs driven by Poisson random measures is investigated.

Let us conclude with introducing some notation used throughout the paper. All random objects are given on a fixed probability space $(\Omega, \mathcal{F}, P)$ equipped with a right-continuous filtration $\mathbb{F} = (\mathcal{F}_t)_{t\ge 0}$, such that $\mathcal{F}_0$ contains all the P-null sets of the complete σ-algebra $\mathcal{F}$. The σ-algebra of predictable subsets of $[0,\infty) \times \Omega$ is denoted by $\mathcal{P}$ and the σ-algebra of Borel subsets of $\mathbb{R}^d$ is denoted by $\mathcal{B}(\mathbb{R}^d)$. The notation $C_0^\infty = C_0^\infty(\mathbb{R}^d)$ stands for the space of real valued smooth functions with compact

support on $\mathbb{R}^d$. For an integer m we use the notation H^m for the Hilbert-Sobolev space $W_2^m(\mathbb{R}^d)$ of generalised functions on $\mathbb{R}^d$. For $m \geq 0$ it is the closure of C_0^∞ in the norm $|\cdot|_m$ defined by

$$|f|_m^2 = \sum_{|\alpha| \leq m} \int_{\mathbb{R}^d} |D^\alpha f(x)|^2 \, dx,$$

where $D^\alpha = D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_d^{\alpha_d}$ for multi-indices $\alpha = (\alpha_1, \alpha_2, \alpha_d) \in \{0, 1, \dots\}^d$, and D^0 is the identity operator. For $m < 0$, H^m is the closure of C_0^∞ in the norm

$$|f|_m = \sup_{g \in C_0^\infty, |g|_{-m} \leq 1} (f, g)_0,$$

where $(f, g)_0$ denotes the inner product in $L_2 = H^0$. We define in the same way the Hilbert-Sobolev space $H^m = H^m(\mathbb{R}^l)$ of $\mathbb{R}^l$-valued functions $g = (g^1, \dots, g^l)$ on $\mathbb{R}^d$, such that $|g|_m^2 = \sum_{k=1}^l |g^k|_m^2$. We use the notation $(\cdot, \cdot)_m$ for the inner product in H^m, and for $m = 0$ we often use the notation $(\cdot, \cdot)$ instead of $(\cdot, \cdot)_0$. For $m \geq 0$ denote by $\langle \cdot, \cdot \rangle_m$ the duality product between H^{m+1} and H^{m-1}, based on the inner product $(\cdot, \cdot)_m$ in H^m. For real numbers A and B we set $A \vee B = \max\{A, B\}$ and $A \wedge B = \min\{A, B\}$. For sequences of random variables $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ the notation $a_n = o(b_n)$ means the existence of a sequence of random variables ξ_n converging almost surely to zero such that almost surely $|a_n| \leq \xi_n |b_n|$ for all n. The notation $a_n = O(b_n)$ means the existence of a finite random variable η such that almost surely $|a_n| \leq \eta |b_n|$ for all n.

2. Formulation of the results

Let $W = (W(t))_{t \in [0,T]}$ be a d_1-dimensional Wiener martingale with respect to $\mathbb{F}$, and consider for every integer $n \geq 1$ an $\mathbb{R}^{d_1}$-valued $\mathcal{F}_t$-adapted continuous process $W_n = (W_n(t))_{t \in [0,T]}$ of finite variation. Define the area processes of W and W_n as

$$A^{ij}(t) := \frac{1}{2} \int_0^t W^i(s) \, dW^j(s) - W^j(s) \, dW^i(s), \quad i, j = 1, 2, ..., d_1, \tag{2.1}$$

$$A_n^{ij}(t) := \frac{1}{2} \int_0^t W_n^i(s) \, dW_n^j(s) - W_n^j(s) \, dW_n^i(s), \quad i, j = 1, 2, ..., d_1, \tag{2.2}$$

and also the process

$$B_n^{ij}(t) := \int_0^t (W^i(s) - W_n^i(s)) \, dW_n^j(s), \qquad i, j = 1, \dots, d_1, \tag{2.3}$$

that will play a crucial role. We denote by $\|q\|(t)$ the first variation of a process q over the interval $[0, t]$ for $t \leq T$.

Let $\gamma > 0$ be a fixed real number and assume that the following conditions hold.

Assumption 2.1. For each $\kappa < \gamma$ almost surely

(i) $\sup_{t \leq T} |W(t) - W_n(t)| = O(n^{-\kappa})$,
(ii) $\sup_{t \leq T} \left|A^{ij}(t) - A_n^{ij}(t)\right| = O(n^{-\kappa}) \quad i \neq j$,
(iii) $||B_n^{ij}||(T) = o(\ln n)$ for all $i, j = 1, \dots, d_1$.

The following remark is shown in [12].

Remark 2.1. Define the matrix-valued process $S_n = (S_n^{ij}(t))$, $t \in [0, T]$ by

$$\begin{aligned} S_n^{ij}(t) &= \int_0^t (W^i(s) - W_n^i(s)) \, dW_n^j(s) - \frac{1}{2} \langle W^i, W^j \rangle(t) \\ &= \int_0^t (W^i(s) - W_n^i(s)) \, dW_n^j(s) - \frac{1}{2} \delta_{ij} t, \qquad i, j = 1, 2, ..., d_1, \end{aligned} \tag{2.4}$$

for each integer $n \geq 1$, where $\langle W^i, W^j \rangle$ denotes the quadratic covariation process of W^i and W^j, and $\delta_{ij} = 1$ for $i = j$ and it is zero otherwise. Then by Itô's formula for $q_n^{ij} := (W^i - W_n^i)(W^j - W_n^j)$ we have

$$S_n^{ij}(t) + S_n^{ji}(t) = q_n^{ij}(0) - q_n^{ij}(t) + R_n^{ij}(t) + R_n^{ji}(t)$$

with

$$R_n^{ij}(t) := \int_0^t (W^i(s) - W_n^i(s))\, dW^j(s).$$

Moreover, given Part (i) in Assumption 2.1, Part (ii) is equivalent to condition (ii′):

$$\sup_{t \leq T} |S_n^{ij}(t)| = O(n^{-\kappa}) \quad \text{(a.s.) for each } \kappa < \gamma, \text{ for } i, j = 1, ..., d_1. \tag{2.5}$$

Assumption 2.1 holds for a large class of approximations W_n of W. The main examples are the following.

Example. (Polygonal approximations) Set $W_n(t) = 0$ for $t \in [0, T/n)$ and

$$W_n(t) = W(t_{k-1}) + n(t - t_k)(W(t_k) - W(t_{k-1}))/T$$

for $t \in [t_k, t_{k+1})$, where $t_k := kT/n$ for integers $k \geq 0$.

Example. (Smoothing) Define

$$W_n(t) = \int_0^1 W(t - u/n)\, du, \quad t \geq 0,$$

where $W(s) := 0$ for $s < 0$.

One can prove, see [13], that these examples satisfy the conditions of Assumptions 2.1 with $\gamma = 1/2$.

Now we formulate the conditions on the operators L_n, M_n^k and their convergence to operators L and M^k. We fix an integer $m \geq 0$ and a real number $K \geq 0$.

Assumption 2.2 (ellipticity). There exists a constant $\lambda \geq 0$ such that for each integer $n \geq 1$ for $dP \times dt \times dx$ almost all $(\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}^d$

$$a_n^{ij}(t, x) z^i z^j \geq \lambda |z|^2,$$

for all $z = (z^1, z^2, ..., z^d) \in \mathbb{R}^d$.

If $\lambda > 0$ then we need the following conditions on the regularity of the coefficients $\mathbf{a}_n = (a_n^{ij}, a_n^i, a_n : i, j = 1, ..., d)$, $\mathbf{b}_n = (b_n^{ik}, b_n^k : i = 1, ..., d; k = 1, ..., d_1)$, $\mathbf{a} := (a^{ij}, a^i, a : i, j = 1, ..., d)$, $\mathbf{b} := (b^{ik}, b^k : i = 1, ..., d; k = 1, ..., d_1)$ for all $n \geq 1$, and on the data u_{n0}, f_n and $g_n = (g_n^k)$, u_0, f, $g = (g^k)$.

Assumption 2.3. The coefficients $\mathbf{a}_n$, $\mathbf{b}_n$ and their derivatives in x up to order $m+4$ are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable functions, and they are in magnitude bounded by K. For each $n \geq 1$, f_n is an H^{m+3}-valued predictable process, $g_n = (g_n^k)$ is an $H^{m+4}(\mathbb{R}^{d_1})$-valued predictable process and u_{n0} is an H^{m+4}-valued $\mathcal{F}_0$-measurable random variable, such that for every $\varepsilon > 0$ almost surely

$$|u_{n0}|_{m+3} = O(n^{\varepsilon}), \quad \int_0^T |f_n|_{m+3}^2\, dt = O(n^{\varepsilon}), \quad \sup_{t \leq T} |g_n(t)|_{m+4} = O(n^{\varepsilon}).$$

One knows, see Theorem 3.2 below, that if Assumption 2.2 with $\lambda > 0$ and Assumption 2.3 hold, then for each $n \geq 1$ there is a unique generalised solution u_n to (1.1)-(1.2).

Assumption 2.4. The coefficients $\mathbf{a}$ and $\mathbf{b}$ and their derivatives in x up to order $m+1$ are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable functions on $\Omega \times H_T$, and they are in magnitude bounded by K. The initial value u_0 is an H^{m+1}-valued $\mathcal{F}_0$-measurable random variable, f is an H^m-valued predictable processes and $g = (g^k)$ is an $H^{m+1}(\mathbb{R}^{d_1})$-valued predictable process such that almost surely

$$\int_0^T |f(t)|_m^2\, dt + \sup_{t \leq T} |g(t)|_{m+1}^2 < \infty.$$

Assumption 2.5. We have

$$\sup_{H_T} |D^\alpha \mathbf{a}_n - D^\alpha \mathbf{a}| = O(n^{-\gamma}), \quad \sup_{H_T} |D^\beta \mathbf{b}_n - D^\beta \mathbf{b}| = O(n^{-\gamma}),$$

for all $|\alpha| \le (m-1) \vee 0$ and $|\beta| \le m+1$, and

$$\int_0^T |f_n(t) - f(t)|_{m-1}^2 \, dt + \sup_{t\le T} |g_n(t) - g(t)|_{m+1}^2 = O(n^{-2\gamma}).$$

Now we formulate our main result when $\lambda > 0$ in Assumption 2.2, and b_n^{ik}, b_n^k and g_n^k do not depend on $t \in [0,T]$.

Theorem 2.2. *Assume that* $\mathbf{b}_n$ *and* g_n *are independent of* t. *Let Assumptions 2.1, 2.2 with* $\lambda > 0$, *2.3, 2.4 and 2.5 hold. Then almost surely*

$$\sup_{t\le T} |u_n(t) - u(t)|_m^2 + \int_0^T |u_n(t) - u(t)|_{m+1}^2 \, dt = O(n^{-2\kappa}), \quad \textit{for any } \kappa < \gamma.$$

In the degenerate case, $\lambda = 0$, instead of Assumptions 2.3, 2.4 and 2.5 we need to impose stronger conditions.

Assumption 2.6. (i) For each $n \ge 1$ there exist functions σ_n^{ir} on $\Omega \times H_T$, for $i = 1, ..., d$ and $r = 1, ..., p$, for some $p \ge 1$, such that $a_n^{ij} = \sigma_n^{ir}\sigma_n^{jr}$ for all $i, j = 1, \dots, d$. (ii) The functions σ_n^{ir}, b_n^i and their derivatives in x up to order $m+6$, the functions a_n^i, a_n, b_n and their derivatives in x up to order $m+5$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions on $\Omega \times H_T$, and in magnitude are bounded by K for $i = 1, ..., d$ and $r = 1, ..., p$. For each $n \ge 1$, f_n is an H^{m+4}-valued predictable process, $g_n = (g_n^k)$ is an $H^{m+5}(\mathbb{R}^{d_1})$-valued predictable process and u_{n0} is an H^{m+4}-valued $\mathcal{F}_0$-measurable random variable, such that for every $\varepsilon > 0$

$$|u_{n0}|_{m+4} = O(n^\varepsilon), \quad \int_0^T |f_n|_{m+4}^2 \, dt = O(n^\varepsilon), \quad \sup_{t\le T} |g_n(t)|_{m+5} = O(n^\varepsilon).$$

Assumption 2.7. The coefficients $\mathbf{a}$ and $\mathbf{b}$ and their derivatives in x up to order $m+2$ are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable functions on $\Omega \times H_T$, and they are in magnitude bounded by K. The initial value u_0 is an H^{m+2}-valued $\mathcal{F}_0$-measurable random variable, f is an H^{m+2}-valued predictable process and $g = (g^k)$ is an $H^{m+3}(\mathbb{R}^{d_1})$-valued predictable process such that

$$\int_0^T |f(t)|_{m+2}^2 \, dt + \sup_{t\le T} |g(t)|_{m+2}^2 < \infty.$$

Assumption 2.8. We have

$$\sup_{H_T} |D^\alpha \mathbf{a}_n - D^\alpha \mathbf{a}| = O(n^{-\gamma}), \quad \sup_{H_T} |D^\beta \mathbf{b}_n - D^\beta \mathbf{b}| = O(n^{-\gamma})$$

for all $|\alpha| \le m$ and $|\beta| \le m+1$, and

$$\int_0^T |f_n - f|_m^2 \, dt = O(n^{-\gamma}), \quad \int_0^T |g_n - g|_{m+1}^2 \, dt = O(n^{-2\gamma}).$$

Remark 2.3. Notice that Assumption 2.6 (i) implies Assumption 2.2 with $\lambda = 0$.

Theorem 2.4. *Assume that* $\mathbf{b}_n$ *and* g_n *do not depend on* t. *Let Assumptions 2.1, 2.6, 2.7 and 2.8 hold. Then*

$$\sup_{t\le T} |u_n - u|_m = O(n^{-\kappa}) \quad \textit{a.s. for each } \kappa < \gamma.$$

Let us now consider the case when all the coefficients and free terms may depend on $t \in [0,T]$. We use the notation

$$h_n := \left(b_n^{ik}, b_n^k, g_n^k : i = 1, ..., d,\ k = 1, ..., d_1\right), \quad n \ge 1,$$
$$h := \left(b^{ik}, b^k, g^k : i = 1, ..., d,\ k = 1, ..., d_1\right).$$

We make the following assumption.

Assumption 2.9. There exist $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable bounded functions

$$\mathbf{b}_n^{(r)} = (b_n^{ik(r)},\, b_n^{k(r)} : i = 1, ..., d,\, k = 1, ..., d_1), \quad r = 0, ..., d_1,\, n \geq 1$$

and $H^0(\mathbb{R}^{d_1})$-valued bounded predictable processes $g_n^{(r)} = (g_n^{k(r)} : k = 1, ..., d_1)$, such that

$$d(h_n(t), \varphi) = (h_n^{(0)}(t), \varphi)\, dt + (h_n^{(k)}(t), \varphi)\, dW_n^k(t), \quad n \geq 1,$$

where $h_n^{(r)} = (\mathbf{b}_n^{(r)}, g_n^{(r)})$. For $r = 0, ...d_1$ and $j = 1, ..., d_1$ there exist $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable bounded functions

$$\mathbf{b}^{(r)} = (b^{ik(r)},\, b^{k(r)} : i = 1, ..., d,\, k = 1, ..., d_1),$$
$$\mathbf{b}^{(jr)} = (b^{ik(jr)},\, b^{k(jr)} : i = 1, ..., d,\, k = 1, ..., d_1),$$

and $H^0(\mathbb{R}^{d_1})$-valued bounded predictable processes $g_n^{k(r)}$ and $g_n^{k(jr)}$, $k = 1, ..., d_1$, such that

$$d(h(t), \varphi) = (h^{(0)}(t), \varphi)\, dt + (h^{(k)}(t), \varphi)\, dW_n^k(t),$$
$$d(h^{(j)}(t), \varphi) = (h^{(j0)}(t), \varphi)\, dt + (h^{(jk)}(t), \varphi)\, dW^k(t)$$

for $\varphi \in C_0^\infty(\mathbb{R}^d)$, for $j = 1, ..., d_1$, where $h^{(r)} = (\mathbf{b}^{(r)}, g^{k(r)} : k = 1, \ldots, d_1)$ and $h^{(jr)} = (\mathbf{b}^{(jr)}, g^{k(jr)} : k = 1, \ldots, d_1)$, $r = 0, ..., d_1$.

If Assumption 2.2 holds with $\lambda > 0$, then we impose the following conditions.

Assumption 2.10. For $n \geq 1$ the coefficients $\mathbf{b}_n^{(r)}$ and their derivatives in x up to order $m + 3$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions on $\Omega \times H_T$, and they are bounded in magnitude by K for $r = 0, 1, ..., d_1$. The functions $g_n^{k(0)}$ are H^{m+2}-valued, $g_n^{k(j)}$ are H^{m+3}-valued predictable processes, such that

$$\int_0^T |g_n^{k(0)}(t)|_{m+2}^2\, dt = O(n^\varepsilon), \quad \sup_{H_T} |g_n^{k(j)}|_{m+3} = O(n^\varepsilon)$$

for each $\varepsilon > 0$ and all $k, j = 1, ..., d_1$.

Assumption 2.11. The coefficients $\mathbf{b}^{(r)}$, $\mathbf{b}^{(jr)}$ and their derivatives in x up to order $m+1$ are $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable functions on $\Omega \times H_T$, and they are bounded in magnitude by K for $r = 0, 1, ..., d_1$ and $j = 1, 2, .., d_1$. The functions $g^{k(r)}$ and $g^{k(jr)}$ are H^{m+1}-valued predictable processes, and are bounded in H^{m+1}, for $r = 0, 1, ..., d_1$ and $j = 1, 2, .., d_1$.

Assumption 2.12. For $j = 1, 2, ..., d_1$ we have

$$\sup_{H_T} |D^\alpha \mathbf{b}_n^{(j)} - D^\alpha \mathbf{b}^{(j)}| = O(n^{-\gamma}) \quad \text{for } |\alpha| \leq m,$$

$$\sup_{t \leq T} |g_n^{k(j)} - g^{k(j)}|_m = O(n^{-\gamma}) \quad \text{for } k = 1, ..., d_1.$$

One knows, see [9], that under the assumptions above the limit u of u_n for $n \to \infty$ exists and satisfies

$$du(t, x) = (Lu(t, x) + f)\, dt + (M^k u(t, x) + g^k) \circ dW^k(t), \quad (t, x) \in H_T \tag{2.6}$$

with initial condition

$$u(0, x) = u_0(x), \quad x \in \mathbb{R}^d, \tag{2.7}$$

where

$$\begin{aligned}(M^k u(t, x) + g^k) \circ\, dW^k =& \tfrac{1}{2}(M^k M^k u(t, x) + M^k g^k(t, x))\, dt \\ &+ (M^k u(t, x) + g^k(t, x))\, dW^k(t) \\ &+ \frac{1}{2} \sum_{k=1}^{d_1} (M^{k(k)} u(t, x) + g^{k(k)}(t, x))\, dt,\end{aligned}$$

with $M^{k(k)} := b^{ik(k)}(t, x) D_i + b^{k(k)}(t, x)$.

We have the following results on the rate of convergence.

Theorem 2.5. *Let Assumptions 2.1, 2.2 with $\lambda > 0$, 2.3, 2.4, 2.5 and 2.9 through 2.12 hold. Then for each $\kappa < \gamma$*

$$\sup_{t\le T}|u_n(t)-u(t)|_m^2+\int_0^T|u_n(t)-u(t)|_{m+1}^2\,dt=O(n^{-2\kappa}),$$

where u is the generalised solution of (2.6)-(2.7).

Let us now consider the case when $\lambda = 0$ in Assumption 2.2.

Assumption 2.13. For $n \geq 1$ the coefficients $\mathbf{b}_n^{(r)}$ and their derivatives in x up to order $m+4$ are $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable functions on $\Omega\times H_T$, and they are bounded in magnitude by K for $r=0,1,...,d_1$. The functions $g_n^{k(0)}$ are H^{m+3}-valued, $g_n^{k(j)}$ are H^{m+4}-valued predictable processes, such that

$$\int_0^T|g^{k(0)}(t)|_{m+3}^2\,dt=O(n^{\varepsilon}),\quad \sup_{H_T}|g^{k(j)}|_{m+4}=O(n^{\varepsilon})$$

for each $\varepsilon>0$ and all $k,j=1,...,d_1$.

Assumption 2.14. The coefficients $\mathbf{b}^{(r)}$, $\mathbf{b}^{(jr)}$ and their derivatives in x up to order $m+2$ are $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable functions on $\Omega\times H_T$, and they are bounded in magnitude by K for $r=0,1,...,d_1$ and $j=1,2,..,d_1$. The functions $g^{k(r)}$ and $g^{k(jr)}$ are H^{m+1}-valued predictable processes, and are bounded in H^{m+1}, for $r=0,1,...,d_1$ and $k,j=1,2,..,d_1$.

Assumption 2.15. For $j=1,2,...,d_1$ we have

$$\sup_{H_T}|D^\alpha\mathbf{b}_n^{(j)}-D^\alpha\mathbf{b}^{(j)}|=O(n^{-\gamma})\quad\text{for } |\alpha|\le m+1,$$

$$\sup_{t\le T}|g_n^{k(j)}-g^{k(j)}|_{m+1}=O(n^{-\gamma})\quad\text{for } k=1,...,d_1.$$

Theorem 2.6. *Let Assumption 2.1, Assumptions 2.6 through 2.9, and Assumptions 2.13 through 2.15 hold. Then*

$$\sup_{t\le T}|u_n-u|_m=O(n^{-\kappa})\quad \textit{for each } \kappa<\gamma,$$

where u is the generalised solution of (2.6)-(2.7).

3. Auxiliaries

3.1. **Existence, uniqueness and known estimates for solutions.** Consider the equation

$$\begin{aligned}du(t,x)=&(\mathcal{L}u(t,x)+f(t,x))\,dt+(\mathcal{M}^ku(t,x)+g^k(t,x))\,dW^k(t)\\ &+(\mathcal{N}^\rho u(t,x)+h^\rho(t,x))\,dB^\rho(t),\quad t\in(0,T],\quad x\in\mathbb{R}^d\end{aligned}\tag{3.1}$$

with initial condition

$$u(0,x)=u_0(x),\quad x\in\mathbb{R}^d,\tag{3.2}$$

where $W=(W^1,...,W^{d_1})$ is a d_1-dimensional Wiener martingale with respect to $(\mathcal{F}_t)_{t\ge0}$, and B^1,...,B^{d_2} are real-valued adapted continuous processes of finite variation over $[0,T]$. The operators $\mathcal{L}$, $\mathcal{M}^k$ and $\mathcal{N}^\rho$ are of the form

$$\mathcal{L}=\mathfrak{a}^{ij}D_iD_j+\mathfrak{a}^iD_i+\mathfrak{a},\quad \mathcal{M}^k=\mathfrak{b}^{ik}D_i+\mathfrak{b}^k,\quad \mathcal{N}^\rho=\mathfrak{c}^{i\rho}D_i+\mathfrak{c}^\rho,$$

where the coefficients $\mathfrak{a}^{ij}$, $\mathfrak{a}^i$, $\mathfrak{a}$, $\mathfrak{b}^{ik}$, $\mathfrak{b}^k$, $\mathfrak{c}^{i\rho}$ and $\mathfrak{c}^\rho$ are $\mathcal{P}\times\mathcal{B}(\mathbb{R}^d)$-measurable real-valued bounded functions defined on $\Omega\times[0,T]\times\mathbb{R}^d$ for all $i,j=1,...,d$, $k=1,...,d_1$ and $\rho=1,...,d_2$. The free terms $f=f(t,\cdot)$, $g^k(t,\cdot)$ and $h^\rho=h^\rho(t,\cdot)$ are H^0-valued predictable processes, and u_0 is an H^1-valued $\mathcal{F}_0$-measurable random variable.

To formulate the notion of the solution we assume that the generalised derivatives in x, D_ja^{ij}, are also bounded functions on $\Omega\times H_T$ for all $i,j=1,...,d$.

Definition 3.1. By a solution of (3.1)-(3.2) we mean an H^1-valued weakly continuous adapted process $u = (u(t))_{t\in[0,T]}$, such that

$$(u(t), \varphi) = (u_0, \varphi) + \int_0^t \{-(\mathfrak{a}^{ij} D_i u, D_j\varphi) + \big((\mathfrak{a}^i - \mathfrak{a}^{ij}_j) D_i u + \mathfrak{a}u + f, \varphi\big)\}\, ds$$
$$+ \int_0^t (\mathcal{M}^k u + g^k, \varphi)\, dW^k + \int_0^t (\mathcal{N}^\rho u + h^\rho, \varphi)\, dB^\rho,$$

holds for $t \in [0, T]$ and $\varphi \in C_0^\infty(\mathbb{R}^d)$, where $\mathfrak{a}^{ij}_j = D_j\mathfrak{a}^{ij}$.

To present those existence and uniqueness theorems from the L_2-theory of stochastic PDEs which we use in this paper, we formulate some assumptions.

Assumption 3.1. There is a constant $\lambda \geq 0$ such that for all $n \geq 1$, $dP \times dt \times dx$ almost all $(\omega, t, x) \in \Omega \times H_T$ we have

$$(\mathfrak{a}^{ij} - \tfrac{1}{2}\mathfrak{b}^{ik}\mathfrak{b}^{jk})z^i z^j \geq \lambda |z|^2 \quad \text{for all } z = (z^1, ..., z^d) \in \mathbb{R}^d.$$

To formulate some further conditions on the smoothness of the coefficients and the data of (3.1)-(3.2) we fix an integer $m \geq 1$. We consider first the case $\lambda > 0$ in Assumption 3.1, and make the following conditions.

Assumption 3.2. The coefficients $\mathfrak{a}^{ij}$, $\mathfrak{a}^i$, $\mathfrak{a}$, $\mathfrak{b}^{ik}$, $\mathfrak{b}^k$, $\mathfrak{c}^{i\rho}$, $\mathfrak{c}^\rho$ and their derivatives in $x \in \mathbb{R}^d$ up to order m are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable real functions on $\Omega \times H_T$ and in magnitude are bounded by K.

Assumption 3.3. The initial value u_0 is an H^m-valued random variable. The free terms $f = f(t)$, $g^k = g^k(t)$, $h^\rho = h^\rho(t)$ are predictable H^m-valued processes such that almost surely

$$\int_0^T |f(t)|^2_{m-1}\, dt < \infty, \quad \int_0^T |g(t)|^2_m\, dt < \infty, \quad \int_0^T |h^\rho(t)|_m\, d\|B^\rho\|(t) < \infty$$

for all $k = 1, ..., d_1$ and $\rho = 1, ..., d_2$, where $|g|_l^2 = \sum_k |g^k|_l^2$ and $|h|_l^2 = \sum_\rho |h^\rho|_l^2$ for ever $l \geq 0$.

Theorem 3.2. *Let Assumptions 3.1 with $\lambda > 0$, 3.2 and 3.3 hold. Then (3.1)-(3.2) has a unique generalised solution u. Moreover, u is an H^m-valued weakly continuous process, it is strongly continuous as an H^{m-1}-valued process, $u(t) \in H^{m+1}$ for $P \times dt$ a.e. (ω, t), and there exist constants $\nu \geq 0$ and $C > 0$ such that for every $l \in [0, m]$*

$$E \sup_{t\leq T} e^{-\nu V}|u|_l^2 + E\int_0^T e^{-\nu V}|u|^2_{l+1}\, dt$$
$$\leq C\left\{|u_0|_l^2 + \int_0^T e^{-\nu V}(|f|^2_{l-1} + |g|_l^2)\, dt + \int_0^T e^{-\nu V}|h^\rho|_l^2\, d\|B^\rho\|\right\},$$

where $V(t) = t + \sum_{\rho=1}^{d_2} \|B^\rho(t)\|$. The constants ν and C depend only on λ, K, d, d_1, d_2 and l.

In the degenerate case, i.e., when $\lambda = 0$ in Assumption 3.1, we need to impose somewhat stronger conditions in the other assumptions of the previous theorem.

Theorem 3.3. *Let Assumptions 3.1 (with $\lambda = 0$), 3.2 and 3.3 hold. Assume, moreover, that the derivatives in $x \in \mathbb{R}^d$ of a^{ij} up to order $m \vee 2$ are bounded real functions on $\Omega \times [0, T] \times \mathbb{R}^d$ for all $i, j = 1, ..., d$, and that $g^k = g^k(t)$ are H^{m+1}-valued predictable processes for all $k = 1, ..., d_1$, such that almost surely*

$$\int_0^T |g(t)|^2_{m+1}\, dt < \infty. < \infty.$$

Then the conclusion of Theorem 3.2 remains valid.

Theorem 3.3 is a slight modification of [11, Theorem 3.1] and can be proved in the same way. We can prove Theorem 3.2 in the same fashion.

3.2. **Inequalities in Sobolev spaces and a Gronwall-type lemma.** In the following lemmas we present some estimate we use in the paper. We consider the differential operators

$$\mathcal{M} = b^i D_i + b^0, \qquad \mathcal{N} = c^i D_i + c^0, \qquad \mathcal{K} = d^i D_i + d^0,$$

and

$$\mathcal{L} = a^{ij} D_i D_j + a^i D_i + a^0,$$

where a^{ij}, a^i, a^0, b^i, b^0, c^i, c^0, d^i and d^0 are Borel functions defined on $\mathbb{R}^d$ for $i, j = 1, \dots, d$. We fix an integer $l \geq 0$ and a constant K. Recall the notation $(\cdot\,,\cdot) = (\cdot\,,\cdot)_0$ for the inner product in $H^0 \equiv L^2(\mathbb{R}^n)$, and $\langle \cdot\,,\cdot \rangle$ for the duality product between H^1 and H^{-1}.

Lemma 3.4. *(i) Assume that b^0 and its derivatives up to order l, and b^i and their derivatives up to order $l \vee 1$ are real functions, in magnitude bounded by K. Then for a constant $C = C(K, l, d)$*

$$|(D^\alpha \mathcal{M} v, D^\alpha v)| \leq C |v|_l^2, \tag{3.3}$$

$$|(D^\alpha \mathcal{M} v, D^\alpha u) + (D^\alpha \mathcal{M} u, D^\alpha v)| \leq C |v|_l |u|_l \tag{3.4}$$

for all $u, v \in H^{l+1}$ and multi-indices α, $|\alpha| \leq l$.

(ii) Assume that b^0, c^0 and their derivatives up to order $l \vee 1$, b^i and c^i and their derivatives up to order $(l+1) \vee 2$ are real functions, in magnitude bounded by K. Then for a constant $C = C(K, l, d)$

$$|(D^\alpha \mathcal{M}\mathcal{N} v, D^\alpha v) + (D^\alpha \mathcal{M} v, D^\alpha \mathcal{N} v)| \leq C |v|_l^2,$$

for all $v \in H^{l+2}$ and multi-indices α, $|\alpha| \leq l$.

(iii) Assume that b^0, c^0, d^0 and their derivatives up to order $(l+1) \vee 2$, b^i, c^i, d^i and their derivatives up to order $(l+2) \vee 3$ are real functions, and in magnitude are bounded by K for $i = 1, ..., d$. Then for a constant $C = C(K, l, d)$

$$\begin{aligned} &|(D^\alpha \mathcal{M}\mathcal{N}\mathcal{K} v, D^\alpha v) + (D^\alpha \mathcal{M}\mathcal{N} v, D^\alpha \mathcal{K} v) \\ &+ (D^\alpha \mathcal{M}\mathcal{K} v, D^\alpha \mathcal{N} v) + (D^\alpha \mathcal{M} v, D^\alpha \mathcal{N}\mathcal{K} v)| \leq C |v|_l^2, \end{aligned}$$

for all $v \in H^{l+3}$ and multi-indices α, $|\alpha| \leq l$.

Proof. These and similar estimates are proved in [9]. For the sake of completeness and the convenience of the reader we present a proof here. We can assume that $v \in C_0^\infty(\mathbb{R}^n)$. Let us start with (i). Integrating by parts, we have

$$(\mathcal{M} D^\alpha v, D^\alpha v) = -(D^\alpha v, \mathcal{M} D^\alpha v) + (D^\alpha v, \bar{m} D^\alpha v),$$

where $\bar{m} := 2b^0 - \sum_{i=1}^d D_i b^i$. Therefore, by writing $[\mathcal{M}, D^\alpha] = D^\alpha \mathcal{M} - \mathcal{M} D^\alpha$,

$$\begin{aligned} (D^\alpha \mathcal{M} v, D^\alpha v) &= (\mathcal{M} D^\alpha v, D^\alpha v) + ([\mathcal{M}, D^\alpha] v, D^\alpha v) \\ &= \tfrac{1}{2}(D^\alpha v, \bar{m} D^\alpha v) + ([\mathcal{M}, D^\alpha] v, D^\alpha v) \leq C |v|_l^2, \end{aligned}$$

by the regularity assumed on the coefficients. Let us write

$$p(v) := (D^\alpha \mathcal{M} v, D^\alpha v) = \tfrac{1}{2}(D^\alpha v, \bar{m} D^\alpha v) + ([\mathcal{M}, D^\alpha] v, D^\alpha v) =: q(v) + r(v).$$

Defining

$$\begin{aligned} 2a(u, v) &:= p(u+v) - p(u) - p(v) = (D^\alpha \mathcal{M} u, D^\alpha v) + (D^\alpha \mathcal{M} v, D^\alpha u), \\ 2b(u, v) &:= q(u+v) - q(u) - q(v) = (\bar{m} D^\alpha u, D^\alpha v), \end{aligned}$$

and

$$2c(u, v) := r(u+v) - r(u) - r(v) = ([\mathcal{M}, D^\alpha] u, D^\alpha v) + ([\mathcal{M}, D^\alpha] v, D^\alpha u),$$

we have

$$a(u, v) = b(u, v) + c(u, v) \tag{3.5}$$

and

$$|a(u, v)| \leq |b(u, v)| + |c(u, v)| \leq C |u|_l |v|_l,$$

which proves the second inequality in (3.3). The identity (3.5) applied with $u = Nv$ establishes that

$$(D^\alpha \mathcal{M}\mathcal{N} v, D^\alpha v) + (D^\alpha \mathcal{M} v, D^\alpha \mathcal{N} v) = (\bar{m} D^\alpha \mathcal{N} v, D^\alpha v)$$

$$+([\mathcal{M}, D^\alpha]\mathcal{N}v, D^\alpha v) + ([\mathcal{M}, D^\alpha]v, D^\alpha \mathcal{N}v).$$

By the previous case,

$$|(D^\alpha \mathcal{M}\mathcal{N}v, D^\alpha v) + (D^\alpha \mathcal{M}v, D^\alpha \mathcal{N}v)| \le C|v|_l^2,$$

and (ii) is proved. For (iii), integrating by parts,

$$\begin{aligned}\widetilde{p}(v) :=&(D^\alpha \mathcal{M}\mathcal{N}v, D^\alpha v) + (D^\alpha \mathcal{M}v, D^\alpha \mathcal{N}v)\\ =&(D^\alpha \mathcal{N}v, \bar{m}D^\alpha v) + ([\mathcal{M}, D^\alpha \mathcal{N}]v, D^\alpha v) + ([\mathcal{M}, D^\alpha]v, D^\alpha \mathcal{N}v)\\ =&\widetilde{q}(v) + \widetilde{r}(v) + \widetilde{s}(v).\end{aligned}$$

By polarizing this last identity as above and letting $u = \mathcal{K}v$, we have

$$\begin{gathered}(D^\alpha \mathcal{M}\mathcal{N}\mathcal{K}v, D^\alpha v) + (D^\alpha \mathcal{M}\mathcal{N}v, D^\alpha \mathcal{K}v)\\ +(D^\alpha \mathcal{M}\mathcal{K}v, D^\alpha \mathcal{N}v) + (D^\alpha \mathcal{M}v, D^\alpha \mathcal{N}\mathcal{K}v)\\ = (D^\alpha \mathcal{N}\mathcal{K}v, \bar{m}D^\alpha v) + (D^\alpha \mathcal{N}v, \bar{m}D^\alpha \mathcal{K}v)\\ +([\mathcal{M}, D^\alpha \mathcal{N}]\mathcal{K}v, D^\alpha v) + ([\mathcal{M}, D^\alpha \mathcal{N}]v, D^\alpha \mathcal{K}v)\\ ([\mathcal{M}, D^\alpha]\mathcal{K}v, D^\alpha \mathcal{N}v) + ([\mathcal{M}, D^\alpha]v, D^\alpha \mathcal{N}\mathcal{K}v) \le C|v|_l^2,\end{gathered}$$

where in the last inequality we used (ii). Hence (iii) is proved. □

Lemma 3.5. *Assume that a^{ij}, b^0 and their derivatives up to order $l \vee 1$, b^i and their derivatives up to order $l \vee 2$, a^i, a^0 and their derivatives up to order l are real functions, in magnitude bounded by K for $i = 1, ..., d$. Then for a constant $C = C(K, l, d)$*

$$|(D^\alpha \mathcal{M}v, D^\alpha \mathcal{L}v) + \langle D^\alpha v, D^\alpha \mathcal{M}\mathcal{L}v\rangle| \le C|v|_{l+1}^2,$$

for $v \in H^{l+2}$ and multi-indices $|\alpha| \le l$.

Proof. Let us check first the case $l = 0$. Denote by $\mathcal{M}^*$ the formal adjoint of $\mathcal{M}$. We have

$$|(\mathcal{M}v, \mathcal{L}v) + \langle v, \mathcal{M}\mathcal{L}v\rangle = |((\mathcal{M} + \mathcal{M}^*)v, \mathcal{L}v)| = |(-b_i^i v + 2b^0 v, \mathcal{L}v)| \le C|v|_1,$$

where $b_i^i = D_i b^i$. For the general case, let $|\alpha| \le l$ and write

$$\begin{aligned}&(D^\alpha \mathcal{M}v, D^\alpha \mathcal{L}v) + \langle D^\alpha v, D^\alpha \mathcal{M}\mathcal{L}v\rangle\\ &= (D^\alpha Mv, D^\alpha Lv) + \langle D^\alpha v, MD^\alpha \mathcal{L}v\rangle + (D^\alpha v, [\mathcal{M}, D^\alpha]\mathcal{L}v)\\ &= (D^\alpha Mv, D^\alpha \mathcal{L}v) + (M^* D^\alpha v, D^\alpha \mathcal{L}v) + (D^\alpha v, [\mathcal{M}, D^\alpha]Lv)\\ &= ([M, D^\alpha]v, D^\alpha \mathcal{L}v) + (-b_i^i D^\alpha v + 2b^0 D^\alpha v, D^\alpha Lv) + (D^\alpha v, [\mathcal{M}, D^\alpha]\mathcal{L}v),\end{aligned}$$

from which the estimate follows. □

The next lemma is a standard fact for elliptic differential operators $\mathcal{L} = a^{ij}D_iD_j + a^iD_i + a^0$

Lemma 3.6. *Assume there exists a constant $\lambda > 0$ such that*

$$a^{ij}(x)z^iz^j \ge \lambda\,|z|^2\,, \quad \text{for all } z, x \in \mathbb{R}^d,$$

and that the derivatives of a^i and a^0 up to order $(l-1) \vee 0$, and the derivatives of a^{ij} up to order $l \vee 1$ are functions, bounded by K, for $i, j = 1, ..., d$. Then there is a constant $C = C(K\lambda, l, d)$ such that

$$(D^\alpha v, D^\alpha \mathcal{L}v) \le C|v|_l^2 - \frac{\lambda}{2}|v|_{l+1}^2,$$

for all $v \in H^{l+2}$ and multi-indices $|\alpha| \le l$.

In the next two lemmas we assume that there exist vector fields

$$\sigma^1 = (\sigma^{i1}(x)), \dots, \sigma^p = (\sigma^{ip}(x)),$$

such that $a^{ij} = \sigma^{ir}\sigma^{jr}$ for all $i, j = 1, \dots, d$. Set

$$\mathcal{N}^r := \sigma^{ir}D_i, \qquad r = 1, \dots, p,$$

and notice that if the σ^r are differentiable then we can write $\mathcal{L} = \sum_{r=1}^p (\mathcal{N}^r)^2 + \mathcal{N}^0$, where $\mathcal{N}^0 = \left(a^j - \sigma^{ir}(D_i\sigma^{jr})\right)D_j + a^0$.

Lemma 3.7. *Assume that the derivatives of σ up to order $(l+1)\vee 2$ and the derivatives of a^i, a^0 up to order $l\vee 1$ are functions, bounded by a constant K for $i=1,...,d$. Then*

$$(D^\alpha \mathcal{L}v, D^\alpha v) \le -\sum_{r=1}^{p} |D^\alpha \mathcal{N}^r v|_0^2 + C|v|_l^2,$$

and

$$|(D^\alpha \mathcal{L}v, D^\alpha u)| \le \sum_{r=1}^{p} |D^\alpha \mathcal{N}^r v|_0^2 + C(|v|_l^2 + |u|_{l+1}^2),$$

for all $v,u\in H^{l+2}$ and multi-indices $|\alpha|\le l$, with a constant $C=C(K,d,l,p)$.

Proof. By Lemma 3.4 (ii) and (iii),

$$\begin{aligned}(D^\alpha v, D^\alpha \mathcal{L}v) =&(D^\alpha v, D^\alpha \mathcal{N}^r\mathcal{N}^r v)_0 + (D^\alpha v, D^\alpha \mathcal{N}^0 v)_0\\ \le& -(D^\alpha \mathcal{N}^r v, D^\alpha \mathcal{N}^r v)_0 + C|v|_l^2,\end{aligned}$$

with a constant $C=C(K,d,l,p)$ and the first inequality of the statement follows. To get the second one we need only note that by interchanging differential operators and by integration by parts we have

$$\begin{aligned}|(D^\alpha \mathcal{N}^r\mathcal{N}^r v, D^\alpha u)| \le&|(D^\alpha \mathcal{N}^r v, \mathcal{N}^r D^\alpha u)| + |([\mathcal{N}^r, D^\alpha]\mathcal{N}^r v, D^\alpha u)| + C|u|_{l+1}^2\\ \le& \sum_{r=1}^{p} |D^\alpha \mathcal{N}^r v|_0^2 + C|u|_{l+1}^2,\\ |(D^\alpha \mathcal{N}^0 v, D^\alpha u)| \le& C(|v|_l^2 + |u|_{l+1}^2),\end{aligned}$$

with constants $C=C(K,d,p,l)$. □

Lemma 3.8. *Assume that the derivatives of σ^i and b^i up to order $(l+2)\vee 3$ and the derivatives of a^i, a^0 and b^0 up to order $(l+1)\vee 2$ are functions, bounded by a constant K for $i=1,...,d$. Then*

$$|(D^\alpha \mathcal{M}\mathcal{L}v, D^\alpha v) + (D^\alpha \mathcal{L}v, D^\alpha \mathcal{M}v)| \le C\sum_{r=1}^{p} |\mathcal{N}^r v|_l^2 + C|v|_l^2,$$

with a constant $C=C(K,l,d,p)$ for all $v\in H^{l+3}$ and multi-indices $|\alpha|\le l$.

Proof. Put $\mathcal{N}=\mathcal{K}=\mathcal{N}^r$, $r=1,\dots,p$ in Lemma 3.4 (iii) and use (i) of the same lemma for $\mathcal{N}^r v$ to get

$$\begin{aligned}&|(D^\alpha \mathcal{M}\mathcal{N}^r\mathcal{N}^r v, D^\alpha v) + (D^\alpha \mathcal{M}v, D^\alpha \mathcal{N}^r\mathcal{N}^r v)|\\ &\le C|v|_l^2 + 2|(D^\alpha \mathcal{M}\mathcal{N}^r v, D^\alpha \mathcal{N}^r v)| \le C|v|_l^2 + C\sum_{r=1}^{p} |\mathcal{N}^r v|_l^2,\end{aligned}$$

and apply Lemma 3.4 (ii) to obtain

$$|(D^\alpha \mathcal{M}\mathcal{N}^0 v, D^\alpha v) + (D^\alpha \mathcal{M}v, D^\alpha \mathcal{N}^0 v)| \le C|v|_l^2,$$

which prove the corollary. □

The following Gronwall type lemma will be useful for our estimates in the next section.

Lemma 3.9. *Let y_n, m_n, Q_n and q_n be sequences of real valued continuous $\mathcal{F}_t$-adapted stochastic processes given on the interval $[0,T]$, such that Q_n is a non-decreasing non-negative process and m_n is a local martingale starting from 0. Let δ, γ be some real numbers with $\delta<\gamma$. Assume that almost surely*

$$0 \le y_n(t) \le \int_0^t y_n(s)\,dQ_n(s) + m_n(t) + q_n(t), \tag{3.6}$$

holds for all $t\in[0,\tau_n]$ and integers $n\ge 1$, where

$$\tau_n = \inf\left\{t\ge 0 : y_n(t) \ge n^{-\delta}\right\} \wedge T.$$

Suppose that almost surely

$$Q_n(\tau_n) = o(\ln n), \quad \sup_{t\le\tau_n} q_n(t) = O(n^{-\gamma}),$$

$$d\langle m_n\rangle \le (y_n^2 + k_n y_n)\, dQ_n, \ \text{on } t\in[0,\tau_n], \quad \int_0^{\tau_n} k_n(s)\, dQ_n = O(n^{-\gamma})$$

for a sequence of non-negative $\mathcal{F}_t$-adapted processes k_n. Then almost surely

$$\sup_{t\le T} y_n(t) = O(n^{-\kappa}), \quad \text{for each } \kappa<\gamma. \tag{3.7}$$

Proof. Let us assume first that $\gamma>0$. The case $\delta=0$ is a slight modification of [12, Lemma 3.8]. It can be proved in the same way by using a suitable generalization of Lemma 3.7 from [12] (see [10]). For $\delta<\gamma\in(0,\infty)$, we see that the conditions of the Lemma are satisfied with $\gamma'=\gamma-\delta$,

$$y_n'(t) = \frac{y_n(t)}{n^{-\delta}}, \quad m_n'(t) = \frac{m_n(t)}{n^{-\delta}}, \quad q_n'(t) = \frac{q_n(t)}{n^{-\delta}}, \quad k_n'(t) = \frac{k_n(t)}{n^{-\delta}}, \tag{3.8}$$

in place of γ, y_n, m_n, q_n and k_n, with $\delta=0$. Hence we have (3.7) for y_n' in place of y_n for each $\kappa<\gamma'$, which gives (3.7) in this case.

Suppose now that $\gamma\le 0$. Take $\bar\gamma\in(\delta,\gamma)$ and set $\gamma':=\gamma-\bar\gamma$, $\delta':=\delta-\bar\gamma$. Define y_n', m_n', q_n' and k_n' as in (3.8) with $\bar\gamma$ in place of δ. Notice that $\gamma'>0$ and that the conditions of the lemma are satisfied by the processes y_n', m_n', q_n' and k_n' in place of y_n, m_n, q_n and k_n, with γ' and δ' instead of γ and δ. Hence (3.7) holds for y_n' and γ' in place of y_n and γ, which gives the lemma. □

Corollary 3.10. *Let σ_n be an increasing sequence of stopping times converging to infinity almost surely. Assume that the conditions of the previous lemma are satisfied with $\bar\tau_n=\inf\{t\ge 0: y_n(t)\ge n^{-\delta}\}\wedge T\wedge\sigma_n$ in place of τ_n. Then its conclusion, (3.7), still holds.*

Proof. The conditions of Lemma 3.9 are satisfied by the processes

$$y_n'(t) = y_n(t\wedge\sigma_n), \quad m_n'(t) = m_n(t\wedge\sigma_n), \quad q_n'(t) = q_n(t\wedge\sigma_n),$$

$$k_n'(t) = k_n(t\wedge\sigma_n),$$

in place of y_n, m_n, q_n and k_n and with $\tau_n'=\inf\left\{t\ge 0: y_n'(t)\ge n^{-\delta}\right\}$ in place of τ_n. Hence

$$\sup_{t\le T} y_n'(t) = O(n^{-\kappa}), \quad \text{for each } \kappa<\gamma. \tag{3.9}$$

Define the set $\Omega_n:=[\sigma_n\ge T]$ and note that since $\sigma_n\nearrow\infty$ almost surely, the set $\Omega'=\cup_{n\ge1}\Omega_n$ has full probability. It remains to prove that the random variable

$$\xi := \sup_{m\ge 1}\sup_{t\le T}\frac{y_m(t)}{m^{-\kappa}}$$

is finite almost surely for all $\kappa<\gamma$. Indeed, take $\omega\in\Omega'$. Then $\omega\in\Omega_n$ for some $n(\omega)\ge 1$, hence $\sigma_m(\omega)\ge T$ for all $m\ge n(\omega)$ and, by (3.9),

$$\sup_{t\le T} y_m(t\wedge\sigma_m) = \sup_{t\le T} y_m(t) \le \zeta_\kappa m^{-\kappa},$$

for all $m\ge n(\omega)$. Since ζ_κ is finite almost surely, so it is ξ. □

4. The growth of the approximations

In this section we estimate solutions u_n of (1.1) for large n. We fix an integer $l\ge 0$, a constant $K\ge 0$, and make the following assumptions.

Assumption 4.1. The derivatives in x of the coefficients a_n^{ij}, a_n^i, a_n, b_n^k up to order $l+1$, and the derivatives in x of b_n^{ik} up to order $(l+1)\vee 2$ are $\mathcal{P}\times\mathcal{B}(\mathbb{R}^d)$-measurable real functions on $\Omega\times[0,T]\times\mathbb{R}^d$ and in magnitude are bounded by K, for all $i,j=1,...,d$, $k=1,...,d_1$, and all $n\ge 1$.

Assumption 4.2. For each $\varepsilon > 0$ almost surely

$$|u_{n0}|_l = O(n^\varepsilon), \quad \int_0^T \left(|f_n|_l^2 + |g_n|_{l+1}^2\right) dt = O(n^\varepsilon), \quad \sup_{t\le T} |g_n(t)|_l^2 = O(n^\varepsilon).$$

We will often use the notation notation $f \cdot V(t)$ for the integral

$$\int_0^t f(s)\, dV(s),$$

when V is a semimartingale and f is a predictable process such that the stochastic integral of f against dV over $[0,t]$ is well-defined. We define

$$\eta_n(t) = \sup_{m\ge n} \sup_{s\le t} |W(s) - W_m(s)| .$$

Notice that Assumption 2.1 (i) clearly implies that $\eta_n(T) = O(n^{-\kappa})$ almost surely for each $\kappa < \gamma$.

First we study the case when b_n^{ik}, b_n^k and g_n^k do not depend on $t \in [0,T]$.

Theorem 4.1. *Assume that b_n^{ik}, b_n^k and g_n^k do not depend on t for $i = 1,...,d$, $k = 1,...,d_1$ and $n \ge 1$. Let Assumptions 2.1* (i) *and* (iii)*, 4.1, 4.2 and 2.2 with $\lambda > 0$ hold. Then for every $\varepsilon > 0$ almost surely*

$$\sup_{t\le T} |u_n(t)|_l^2 + \int_0^T |u_n|_{l+1}^2 \, dt = O(n^\varepsilon).$$

Proof. Assume for the moment that $u_{n0} \in H^{l+1}$ almost surely. Recall that we are assuming that W_n^k is of bounded variation, $n \ge 1$, $k = 1, \dots, d_1$. Then by Theorem 3.2, under Assumptions 4.1, 4.2 and 2.2 with $\lambda > 0$ there is a unique generalised solution u_n of (1.1)-(1.2), and it is an H^{l+1}-valued weakly continuous process such that almost surely

$$\int_0^T |u_n(t)|_{l+2}^2 \, dt < \infty.$$

In particular,

$$\begin{aligned}(u_n(t), \varphi)_0 &= (u_{n0}, \varphi)_0 \\ &+ \int_0^t \big[-(a_n^{ij} D_i u_n, D_j \varphi)_0 + \big((a_n^i - D_j a_n^{ij}) D_i u_n + a_n u_n + f_n, \varphi\big)_0 \big]\, ds \\ &+ \int_0^t (b_n^{ik} D_i u_n + b_n^k u_n + g_n^k, \varphi)_0 \, dW_n^k\end{aligned}$$

holds for all $t \in [0,T]$ and $\varphi \in C_0^\infty(\mathbb{R}^d)$. Substituting here $\sum_{|\alpha|\le l} (-1)^{|\alpha|} D^{2\alpha}\varphi$ in place of φ and integrating by parts, we get

$$\begin{aligned}(u_n(t), \varphi)_l &= (u_{n0}, \varphi)_l \\ &+ \int_0^t \big[-(a_n^{ij} D_i u_n, D_j \varphi)_l + \big((a_n^i - D_j a_n^{ij}) D_i u_n + a_n u_n + f_n, \varphi\big)_l \big]\, ds \\ &+ \int_0^t (b_n^{ik} D_i u + b_n^k u_n + g_n^k, \varphi)_l \, dW_n^k(s).\end{aligned}$$

Hence using Itô's formula in the triple $H^{m+1} \hookrightarrow H^m \equiv (H^m)^* \hookrightarrow H^{m-1}$, we have

$$\begin{aligned}d|u_n|_l^2 = 2(u_n, L_n u_n + f_n)_l \, dt + 2(u_n, M_n^k u_n + g_n^k)_l \, dW_n^k = 2(u_n, L_n u_n + f_n)_l \, dt \\ +2(u_n, M_n^k u_n + g_n^k)_l \, dW^k + 2(u_n, M_n^k u_n + g_n^k)_l \, d(W_n^k - W^k).\end{aligned}$$

Hence, integrating by parts in the last term above we have

$$|u_n|_l^2 = \sum_{j=1}^6 I_n^{(j)}, \tag{4.1}$$

where

$$
\begin{aligned}
I_n^{(1)} &= |u_{n0}|_l^2\,,\\
I_n^{(2)} &= 2(u_n, L_n u_n + f_n)_l \cdot t,\\
I_n^{(3)} &= 2(u_n, M_n^k u_n + g_n^k)_l \cdot W^k(t),\\
I_n^{(4)} &= 2(u_n, M_n^k u_n + g_n^k)_l (W_n^k - W^k)\Big|_0^t,\\
I_n^{(5)} &= 2\left(\{(L_n u_n + f_n, M_n^k u_n + g_n^k)_l + \langle u_n, M_n^k(L_n u_n + f_n)\rangle_l\}(W^k - W_n^k)\right) \cdot t,\\
I_n^{(6)} &= 2\left((M_n^j u_n + g^j, M_n^k u_n + g_n^k)_l + \langle u_n, M_n^k(M_n^j u_n + g_n^j)\rangle_l\right) \cdot B_n^{kj}(t),
\end{aligned}
$$

and $\langle\cdot\,,\cdot\rangle_l$ is the duality product between H^{l+1} and H^{l-1}, based on the inner product $(\cdot\,,\cdot)_l$ in H^l. Using Lemma 3.6 we obtain

$$
I_n^{(2)} \le \left[C\,|u_n|_l^2 - \lambda |u_n|_{l+1}^2 + C|f_n|_{l-1}^2\right] \cdot t
$$

with a constant $C = C(K, l, d, \lambda)$. The term $I_n^{(3)}$ is a continuous local martingale starting from 0, such that its quadratic variation, $\langle I_n^{(3)}\rangle$, satisfies, by Lemma 3.4 *(i)*

$$
d\langle I_n^{(3)}\rangle \le C(|u_n|_l^4 + |u_n|_l^2 |g_n|_l^2)\,dt, \tag{4.2}
$$

and also by Lemma 3.4 *(i)* we have

$$
I_n^{(4)}(t) \le C \sup_{0\le s\le t}\left(|u_n(s)|_l^2 + |g_n(s)|_l^2\right)\eta_n(t)
$$

with a constant $C = C(K, d, d_1, l)$, where $|g_n|_l^2 = \sum_k |g_n^k|_l^2$. Using Lemmas 3.4 *(ii)* and 3.5, we have

$$
I_n^{(5)}(t) \le C\left[|u_n|_{l+1}^2 + |f_n|_l^2 + |g_n|_{l+1}^2\right]\eta_n \cdot t,
$$

and

$$
I_n^{(6)}(t) \le C\left[|u_n|_l^2 + |g_n|_l^2\right] \cdot \|B_n\|(t),
$$

with $\|B_n\| := \sum_{j,k} \|B_n^{jk}\|$ and a constant $C = C(K, d, d_1, l)$. Therefore, from (4.1),

$$
\begin{aligned}
|u_n(t)|_l^2 \le{}& \int_0^t (|u_n(s)|_l^2 + |g_n|_l^2)\,dQ_n(s) + m_n(t)\\
&- \int_0^t (\lambda - C\eta_n(s))\,|u_n(s)|_{l+1}^2\,ds + C\int_0^t (|f_n(s)|_l + |g_n|_{l+1}^2)(1+\eta_n(s))\,ds\\
&+ |u_0|_l^2 + C\sup_{0\le s\le t}\left(|u_n(s)|_l^2 + |g_n(s)|_l^2\right)\eta_n(t),
\end{aligned} \tag{4.3}
$$

with $m_n = I_n^{(3)}$ and $Q_n(s) = C(s + \|B_n\|(s))$. Define

$$
\sigma_n := \inf\{t \ge 0 : 2C\eta_n(t) \ge \lambda\}\,,
$$

and note that almost surely

$$
\begin{aligned}
y_n(t) := |u_n(t)|_l^2 + \frac{\lambda}{2}\int_0^t |u_n(s)|_{l+1}^2\,ds &\le \int_0^t |u_n(s)|_l^2\,dQ_n(s) + m_n(t) + q_n(t)\\
&\le \int_0^t y_n(s)\,dQ_n(s) + m_n(t) + q_n(t) \quad \text{for all } t \in [0, \sigma_n],
\end{aligned} \tag{4.4}
$$

with $q_n = q_n^{(1)} + q_n^{(2)}$, where

$$\begin{aligned} q_n^{(1)}(t) =& |u_{n0}|_l^2 + C\eta_n(t) \sup_{0\le s\le t} |g_n|_l^2 \\ &+ C\int_0^t (1+\eta_n)(|f_n(s)|_l^2 + |g_n|_{l+1}^2)\,ds + C\int_0^t |g_n|_l^2\,d\|B_n\|(s), \\ q_n^{(2)}(t) =& C\eta_n(t) \sup_{0\le s\le t} |u_n(s)|_l^2. \end{aligned}$$

Due to Assumptions 2.1 (i)&(iii) and 4.2 we have

$$\sup_{t\le T} |q_n^{(1)}(t)| = O(n^{\varepsilon}) \quad \text{almost surely for each } \varepsilon > 0.$$

For a given $\kappa \in (0,\gamma)$ and any $\varepsilon \in (0,\kappa)$ take $\bar\varepsilon \in (\varepsilon,\kappa)$ and define

$$\tau_n = \inf\{t \ge 0 : |u_n(t)|_l^2 \ge n^{\bar\varepsilon}\}.$$

Then clearly

$$\sup_{t\le \tau_n} |q_n^{(2)}(t)| = O(n^{\varepsilon}) \quad \text{a.s. for each } \varepsilon > 0.$$

Thus

$$\sup_{t\le \tau_n} |q_n(t)| = O(n^{\varepsilon}) \quad \text{a.s. for each } \varepsilon > 0.$$

Now taking into account (4.2) and noting that $\sigma_n \nearrow \infty$, we finish the proof of the theorem by applying Corollary 3.10 to (4.4). Since our estimates do not depend on the norm of u_{n0} in H^{l+1} but on its norm in H^l, by a standard approximation argument we can relax the assumption that u_{n0} is almost surely in H^{l+1}. □

In the case when b_n^{ik}, b_n^k and g_n^k depend on t we make the following assumption.

Assumption 4.3. For each $n \ge 1$, $i = 1,2,...,d$ and $k = 1,...,d_1$ there exist real-valued $\mathcal{P}\times\mathcal{B}(\mathbb{R}^d)$-measurable functions $b_n^{ik(r)}$, $b_n^{k(r)}$ on $\Omega\times[0,T]\times\mathbb{R}^d$ and H^0-valued predictable processes $g_n^{k(r)}$ for $r = 0,1,...,d_1$, such that almost surely

$$\begin{aligned} d(b_n^{ik}(t),\varphi) &= (b_n^{ik(0)}(t),\varphi)\,dt + (b_n^{ik(p)}(t),\varphi)\,dW_n^p(t), \\ d(b_n^{k}(t),\varphi) &= (b_n^{k(0)}(t),\varphi)\,dt + (b_n^{k(p)}(t),\varphi)\,dW_n^p(t), \\ d(g_n^{k}(t),\varphi) &= (g_n^{k(0)}(t),\varphi)\,dt + (g_n^{k(p)}(t),\varphi)\,dW_n^p(t), \end{aligned}$$

for all $i = 1,....,d$, $k = 1,...,d_1$, every $n \ge 1$ and $\varphi \in C_0^\infty(\mathbb{R}^d)$. The functions $b_n^{ik(r)}$ together with their derivatives in x up to order $l\vee 1$ and the functions $b_n^{k(r)}$ together with their derivatives in x up to order l are $\mathcal{P}\times\mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K for all $n \ge 1$ and $r = 0,1,...,d_1$. Moreover, for each $\varepsilon > 0$

$$\int_0^T |g_n^{k(0)}|_{l-1}^2\,dt = O(n^{\varepsilon}), \quad \sup_{t\le T}\sum_{p=1}^d |g_n^{k(p)}|_l^2 = O(n^{\varepsilon}) \quad \text{for } k = 1,...,d_1.$$

Theorem 4.2. *Let the assumptions of Theorem 4.1 together with Assumption 4.3 hold. Then we have the conclusion of Theorem 4.1.*

Proof. We can follow the proof of the previous theorem with minor changes. We need only add an additional term,

$$I_n^{(7)} = 2\{(W^k - W_n^k)(u_n, M_n^{k(0)}u_n + g^{k(0)})_l\}\cdot t + 2(u_n, M_n^{k(p)}u_n + g^{k(p)})_l \cdot B_n^{kp}(t),$$

to the right-hand side (4.1), where for each $n \ge 1$,

$$M_n^{k(r)} = b_n^{ik(r)}D_i + b_n^{k(r)},$$

for $k = 1,...,d_1$ and $r = 0,..,d_1$. Clearly, $2(u_n, g^{k(0)})_l \le |u_n|_{l+1}^2 + |g^k|_{l-1}^2$, and hence

$$|2(W^k - W_n^k)(u_n, M_n^{k(0)}u_n + g^{k(0)})_l|$$

$$\leq \eta_n(|u_n|_{l+1}^2 + |g_n^{(0)}|_{l-1}^2 + C|u_n|_l^2),$$

and, by Lemma 3.4 *(i)*,

$$2|(u_n, M_n^{k(p)} u_n + g^{k(p)})_l| \leq C|u_n|_l^2 + |g^{k(p)}|_l^2$$

with a constant $C = C(K, d, d_1, l)$, where $|g_n^{(0)}|_{l-1}^2 = \sum_{k=1}^{d_1} |g_n^{k(0)}|_{l-1}^2$. Thus inequality (4.3) holds with the additional term

$$q_n^{(3)}(t) = \eta_n(t) \int_0^t |g_n^{(0)}(s)|_{l-1}^2 \, ds + \int_0^t |g^{k(p)}(s)|_l^2 \, d\|B^{kp}\|(s)$$

added to its left-hand side and with a constant $C = C(\lambda, K, d_1, l)$. Since due to Assumptions 4.3 and 2.1 (iii), for each $\varepsilon > 0$ we have

$$\sup_{t \leq T} q_n^{(3)}(t) = O(n^\varepsilon) \quad \text{almost surely},$$

we can finish the proof as in the proof of Theorem 4.1. □

Let us consider now the degenerate case, $\lambda = 0$ in Assumption 2.2.

Assumption 4.4. For each $n \geq 1$ there exist real-valued functions σ_n^{ip} on $\Omega \times H_T$ for $p = 1, 2, ..., d_2$ such that $a_n^{ij} = \sigma_n^{ip}\sigma_n^{jp}$ for all $i, j = 1, \ldots, d$. For all $n \geq 1$ the functions σ_n^{ip} and b_n^i and their derivatives in $x \in \mathbb{R}^d$ up to order $(l+2) \vee 3$, the functions a_n^i, a_n, b_n and their derivatives in x up to order $(l+1) \vee 2$ are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K, for all $i = 1, ..., d$ and $p = 1, ..., d_2$.

Assumption 4.5. Let Assumption 4.3 hold and assume that for each $\varepsilon > 0$

$$\int_0^T |g_n^{k(0)}(t)|_l^2 \, dt = O(n^\varepsilon) \quad \text{almost surely for each } k = 1, 2, ..., d_1.$$

Theorem 4.3. *Let Assumptions 2.1* (i)&(iii)*, 2.2 (with $\lambda = 0$), 4.2, 4.4 and 4.5 hold. Then*

$$\sup_{t \leq T} |u_n(t)|_l^2 + \sum_{r=1}^{d_2} \int_0^T |N_n^r u_n|_l^2 \, ds = O(n^\varepsilon), \quad \textit{for every } \varepsilon > 0 \textit{ almost surely},$$

where $N_n^r = \sigma_n^{ir} D_i$, for $r = 1, \ldots, d_2$.

Proof. The proof follows the lines of that of Theorems 4.1 and 4.2, but instead of estimating $I_n^{(2)}$ and $I_n^{(5)}$ separately, we estimate their sum as follows. Note that

$$\begin{aligned} dI_n^{(5)} =& 2\{(L_n u_n + f_n, M_n^k u_n + g_n^k)_l + \langle u_n, M_n^k(L_n u_n + f_n)\rangle_l\}(W^k - W_n^k)\, dt \\ =& 2\{(L_n u_n, M_n^k u_n)_l + \langle u_n, M_n^k L_n u_n\rangle_l + (L_n u_n, g_n^k)_l\}(W^k - W_n^k)\, dt \\ &+ 2\{(f_n, M_n^k u_n)_l + (u_n, M_n^k f_n)_l\}(W^k - W_n^k)\, dt \\ &+ 2(f_n, g_n^k)_l(W^k - W_n^k)\, dt. \end{aligned}$$

Using Lemmas 3.8, 3.7, 3.4 *(i)* and *(ii)*, we obtain

$$|(L_n u_n, M_n^k u_n)_l + \langle u_n, M_n^k L_n u_n\rangle_l| \leq C \sum_{r=1}^{d_2} |N_n^r u_n|_l^2 + C|u_n|_l^2,$$

$$(u_n, L_n u_n + f_n)_l \leq -\sum_{r=1}^{d_2} |N_n^r u_n|_l^2 + C|u_n|_l^2 + |f_n|_l^2,$$

$$|(f_n, M_n^k u_n)_l + (u_n, M_n^k f_n)_l| \leq C\left(|u_n|_l^2 + |f_n|_l^2\right),$$

$$|(L_n u_n, g_n^k)_l| \leq \sum_{r=1}^{d_2} |N_n^r u_n|^2 + C(|u_n|_l^2 + |g_n^k|_{l+1}^2)$$

with a constant $C = C(K, l, d, d_2)$. Hence

$$dI_n^{(5)}(t) \le C\eta_n(t) \left(\sum_{r=1}^{d_2} |N_n^r u_n|_l^2 + |g_n|_{l+1}^2 + |f_n|_l^2 + |u_n|_l^2 \right) dt,$$

and recalling that $I_n^{(2)}(t) = 2(u_n, L_n u_n + f_n)_l \cdot t$, we get

$$\begin{aligned} I_n^{(2)}(t) + I_n^{(5)}(t) \le & C\left\{(\eta_n + 1)(|u_n|_l^2 + |f_n|_l^2) + \eta_n \left|g_n\right|_{l+1}^2)\right\} \cdot t \\ & + \{(C\eta_n - 2) \sum_{r=1}^{d_2} |N_n^r u_n|_l^2\} \cdot t. \end{aligned}$$

To estimate $I_n^{(7)}$ we use that, by Lemma 3.4*(i)*,

$$|2(W^k - W_n^k)\{(u_n, M_n^{k(0)} u_n + g_n^{k(0)})_l\}| \le C\eta_n(|u_n|_l^2 + |g_n^{(0)}|_l^2)$$

with a constant C and $|g_n^{(0)}|_l^2 = \sum_k |g_n^{k(0)}|_l^2$. Thus using the estimates for $I_n^{(1)}$, $I_n^{(3)}$, $I_n^{(4)}$ and $I_n^{(6)}$ given in the proof of Theorem 4.1, and defining

$$\sigma_n = \inf\{t \ge 0 : C\eta_n \ge 1\},$$

and

$$y_n(t) = |u_n(t)|_l^2 + \sum_{r=1}^{d_2} \int_0^t |K_r u_n|_l^2 \, ds,$$

we get

$$y_n(t) \le \int_0^t y_n(s)\, dQ_n(s) + m_n(t) + q_n(t) \quad \text{almost surely for all } t \in [0, \sigma_n],$$

with

$$Q_n(s) = C\{(\eta_n + 1)s + \|B_n\|(s) + \eta_n\}, \quad m_n = I_n^{(3)}, \quad q_n = q_n^{(1)} + q_n^{(2)} + \bar{q}_n^{(3)},$$

$$\bar{q}_n^{(3)} := \int_0^t |g_n^{(0)}(s)|_l^2 \sup_{r \le t} |W(r) - W_n(r)|\, ds + \int_0^t |g_n^{k(p)}(s)|_l^2 \, d\|B_n^{kp}\|(s),$$

where $\|B_n\|(s) = \sum_{k,p} \|B_n^{kp}\|(s)$, $q_n^{(1)}$ and $q_n^{(2)}$ are defined in the proof Theorem 4.1, and C is a constant depending only on K, d, d_1, d_2 and l. Hence the proof is the same as that of Theorem 4.1. □

5. Rate of convergence results for SPDEs

Here we present two theorems on rate of convergence which provide us with a technical tool to prove our main results. Consider for each integer $n \ge 1$ the problem

$$\begin{aligned} du_n(t,x) =& (\mathcal{L}_n u_n(t,x) + f_n(t,x))\, dt + (\mathcal{M}_n^k u_n(t,x) + g_n^k(t,x))\, dW^k(t) \\ & + (\mathcal{N}_n^\rho u_n(t,x) + h_n^\rho(t,x))\, dB_n^\rho(t), \quad (t,x) \in H_T, \end{aligned} \tag{5.1}$$

$$u_n(0,x) = u_{n0}(x) \quad x \in \mathbb{R}^d, \tag{5.2}$$

where $B_n = (B_n^\rho)$ is an $\mathbb{R}^{d_2}$-valued continuous adapted process of finite variation on $[0,T]$. The operators $\mathcal{L}_n$, $\mathcal{M}_n^k$ and $\mathcal{N}_n^\rho$ are of the form

$$\mathcal{L}_n = \mathfrak{a}_n^{ij}(t,x) D_{ij} + \mathfrak{a}_n^i(t,x) D_i + \mathfrak{a}_n(t,x),$$

$$\mathcal{M}_n^k = \mathfrak{b}_n^{ik}(t,x) D_i + \mathfrak{b}_n^k(t,x), \quad \mathcal{N}_n^\rho = \mathfrak{c}_n^{i\rho}(t,x) D_i + \mathfrak{c}_n^\rho(t,x) \tag{5.3}$$

where $\mathfrak{a}_n^{ij}$, $\mathfrak{a}_n^i$, $\mathfrak{a}_n$, $\mathfrak{b}_n^{ik}$, $\mathfrak{b}_n^k$, $\mathfrak{c}_n^{i\rho}$ and $\mathfrak{c}_n^\rho$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable real functions on $\Omega \times H_T$ for $i, j = 1, ..., d$, $k = 1, ..., d_1$, $\rho = 1, ..., d_2$ and $n \ge 1$. For each n the initial value u_{n0} is an H^0-valued $\mathcal{F}_0$-measurable random variable, f_n is an H^{-1}-valued predictable process and g_n^k and h_n^ρ are H^0-valued predictable processes for $k = 1, ..., d_1$ and $\rho = 1, ..., d_2$. We use the notation $|g_n|_r^2 = \sum_k |g_n^k|_r^2$ and $|h_n|_r^2 = \sum_\rho |h_n^\rho|_r^2$ for $r \ge 0$.

Let $l \ge 0$ be an integer and let $K \ge 0$, $\gamma > 0$ be fixed constants.

We assume the *stochastic parabolicity condition.*

Assumption 5.1. There is a constant $\lambda \geq 0$ such that for all $n \geq 1$, $dP \times dt \times dx$ almost all $(\omega, t, x) \in \Omega \times H_T$ we have

$$(\mathfrak{a}_n^{ij} - \tfrac{1}{2}\mathfrak{b}_n^{ik}\mathfrak{b}_n^{jk})z^i z^j \geq \lambda |z|^2 \quad \text{for all } z = (z^1, ..., z^d) \in \mathbb{R}^d.$$

In the case when $\lambda > 0$ we will use the the following conditions.

Assumption 5.2. The coefficients $\mathfrak{a}_n^{ij}$, $\mathfrak{b}_n^{ik}$, $\mathfrak{c}_n^{i\rho}$ and their derivatives in x up to order $l \vee 1$ are bounded in magnitude by K for all $i, j = 1, ..., d$, $k = 1, ..., d_1$, $\rho = 1, ..., d_2$ and $n \geq 1$. The coefficients $\mathfrak{a}_n^i$, $\mathfrak{a}_n$, $\mathfrak{b}_n^k$, $\mathfrak{c}_n^\rho$ and their derivatives in x up to order l are bounded in magnitude by K for all $i = 1, ..., d$, $k = 1, ..., d_1$, $\rho = 1, ..., d_2$ and $n \geq 1$.

Assumption 5.3. We have $|u_{n0}|_l = O(n^{-\gamma})$,

$$\int_0^T |f_n(s)|_{l-1}^2 \, ds = O(n^{-2\gamma}), \quad \int_0^T |g_n(s)|_l^2 \, ds = O(n^{-2\gamma}),$$

$$\int_0^T |h_n^\rho(t)|_l^2 d\|B^\rho\|(t) = O(n^{-2\gamma}), \quad \sum_{\rho=1}^{d_2} \|B_n^\rho\|(T) = o(\ln n).$$

Let u_n be a generalised solution of (5.1)-(5.2) in the sense of Definition 3.1, such that u_n is an H^l-valued weakly continuous process, $u_n(t) \in H^{l+1}$ for $P \times dt$-almost every $(\omega, t) \in \Omega \times [0, T]$, and almost surely

$$\int_0^T |u_n(t)|_{l+1}^2 \, dt < \infty.$$

Then for $n \to \infty$ we have the following result.

Theorem 5.1. *Let Assumptions 5.2, 5.3 and 5.1 with $\lambda > 0$ hold. Then*

$$\sup_{t \leq T} |u_n(t)|_l^2 + \int_0^T |u_n(t)|_{l+1}^2 \, dt = O(n^{-2\kappa}) \quad \textit{a.s. for } \kappa < \gamma. \tag{5.4}$$

Proof. By the definition of the generalised solution

$$(u_n(t), \varphi) = (u_{n0}, \varphi)$$
$$+ \int_0^t \big[-(\mathfrak{a}_n^{ij} D_i u_n(s), D_j \varphi) + \big((\mathfrak{a}_n^i - (D_j \mathfrak{a}_n^{ij})) D_i u_n(s) + \mathfrak{a}_n u_n(s) + f_n(s), \varphi\big)\big] \, ds$$
$$+ \int_0^t (\mathcal{M}^k u_n(s) + g_n^k(s), \varphi) \, dW^k(s) + \int_0^t (\mathcal{N}_n^\rho u_n(s) + h_n^\rho(s), \varphi) \, dB_n^\rho(s)$$

for all $\varphi \in C_0^\infty(\mathbb{R}^d)$. By Itô's formula

$$|u_n(t)|_0^2 = |u_{n0}|_0^2 + \int_0^t \mathcal{I}_n(s) \, ds + \int_0^t \mathcal{J}_n^k(s) \, dW^k(s) + \int_0^t \mathcal{K}_n^\rho(s) \, dB^\rho(s),$$

with

$$\mathcal{I}_n = -2(\mathfrak{a}_n^{ij} D_i u_n, D_j u_n) + 2\big((\mathfrak{a}_n^i - \mathfrak{a}_{nj}^{ij}) D_i u_n + \mathfrak{a}_n u_n + f_n, u_n\big)$$
$$+ (\mathfrak{b}_n^{ik} D_i u_n + \mathfrak{b}_n^k u_n + g_n^k, \mathfrak{b}_n^{jk} D_j u_n + \mathfrak{b}_n^k u_n + g_n^k),$$
$$\mathcal{J}_n^k = 2(\mathcal{M}_n^k u_n + g_n^k, u_n), \quad \mathcal{K}_n^\rho = 2(\mathcal{N}_n^\rho u_n + h_n^\rho, u_n),$$

where $\mathfrak{a}_{nj}^{ij} := D_j \mathfrak{a}_n^{ij}$. By Assumption 5.1

$$-2(\mathfrak{a}_n^{ij} D_i u_n, D_j u_n) + (\mathfrak{b}_n^{ik} D_i u_n, \mathfrak{b}_n^{jk} D_j u_n) \leq -2\lambda \sum_{i=1}^d |D_i u_n|^2.$$

Hence by standard estimates and Lemma 3.4 *(i)*,

$$\mathcal{I}_n \leq -\lambda |u_n|_1^2 + C(|u_n|^2 + |f_n|_{-1}^2 + |g_n|^2), \quad |\mathcal{K}_n^\rho| \leq C(|u_n|^2 + |h_n^\rho|^2),$$
$$|\mathcal{J}_n^k|^2 \leq C(|u_n|^4 + |g_n^k|^2 |u_n|^2),$$

with a constant $C = C(K, \lambda, d, d_1)$. Consequently, almost surely

$$|u_n(t)|^2 + \lambda \int_0^t |u_n|_1^2\, ds \le |u_{n0}|^2 + \int_0^t |u_n|^2\, dQ_n$$

$$+ \int_0^t \mathcal{J}_n^k\, dW^k + C\int_0^t \left(|f_n|_{-1}^2 + |g_n|^2\right)\, ds + C\int_0^t |h_n^\rho|^2\, d\|B_n^\rho\| \tag{5.5}$$

for all $t \in [0,T]$ and $n \ge 1$, where $Q_n(s) = C\left(s + \sum_{\rho=1}^{d_2} \|B_n^\rho\|(s)\right)$. By Assumption 5.3 we have

$$Q_n(T) = o(\ln n), \quad |u_{n0}|^2 + \int_0^T |f_n|_{-1}^2\, ds + \int_0^T |h_n^\rho(s)|^2\, d\|B_n^\rho\|(s) = O(n^{-2\gamma})$$

almost surely. Notice also that for

$$m_n(t) := \int_0^t \mathcal{J}_n^k\, dW^k(s) \tag{5.6}$$

we have

$$d\langle m_n\rangle = \sum_{k=1}^{d_1} |\mathcal{J}_n^k|^2\, dt \le Cd_1(|u_n(t)|^4 + \gamma_n|g_n|^2|u_n(t)|^2)\, dQ_n,$$

where $\gamma_n(t) = dt/dQ_n(t)$. Due to Assumption 5.3

$$\int_0^T \gamma_n|g_n|^2\, dQ_n = \int_0^T |g_n|^2\, ds = O(n^{-2\gamma}) \quad \text{almost surely.}$$

Hence applying Lemma 3.9 with

$$y_n(t) := |u_n|^2 + \lambda\int_0^t |u_n|_1^2\, ds,$$

$$q_n(t) := |u_{n0}|^2 + C\int_0^t (|f_n|_{-1}^2 + |g_n|^2)\, ds + C\int_0^t |h_n^\rho|^2\, d\|B_n^\rho\|,$$

and with m_n defined in (5.6), from (5.5) we get (5.4) for $l = 0$.

Assume now that $l \ge 1$ and let α be a multi-index such that $1 \le |\alpha| \le l$. Then $\alpha = \beta + \gamma$ for some multi-index γ of length 1, and by definition of the generalised solution we get

$$(D^\alpha u_n(t), \varphi) = (D^\alpha u_{n0}, \varphi)$$

$$-\int_0^t \left[(D^\alpha \mathfrak{a}_n^{ij} D_i u_n, D_j\varphi) + \left(D^\beta\{(\mathfrak{a}_n^i - \mathfrak{a}_{nj}^{ij})D_i u_n + \mathfrak{a}_n u + f_n\}, D^\gamma\varphi\right)\right] ds$$

$$+\int_0^t (D^\alpha \mathcal{M}^k u_n + D^\alpha g^k, \varphi)\, dW^k(s) + \int_0^t (D^\alpha \mathcal{N}_n^\rho u_n + D^\alpha h_n^\rho, \varphi)\, dB^\rho(s).$$

Hence by Itô's formula

$$|D^\alpha u_n(t)|_0^2 = |D^\alpha u_{n0}|_0^2 + \int_0^t \mathcal{I}_n^\alpha\, ds + \int_0^t \mathcal{K}_n^{\rho\alpha}\, dB^\rho + m_n^\alpha(t),$$

with

$$m_n^\alpha(t) = \int_0^t \mathcal{J}_n^{k\alpha}\, dW^k, \quad \mathcal{J}_n^{k\alpha} = 2(D^\alpha\mathcal{M}_n^k u_n + D^\alpha g_n^k, D^\alpha u_n),$$

$$\mathcal{I}_n^\alpha = -2(D^\alpha \mathfrak{a}_n^{ij} D_i u_n, D_j D^\alpha u_n) - 2\left(D^\beta\{(\mathfrak{a}_n^i - \mathfrak{a}_{nj}^{ij})D_i u_n + \mathfrak{a}_n u + f_n\}, D^\alpha D^\gamma u_n\right)$$

$$+(D^\alpha\{\mathfrak{b}_n^{ik} D_i u_n + \mathfrak{b}^k u_n + g^k\}, D^\alpha\{\mathfrak{b}_n^{jk} D_j u_n + \mathfrak{b}_n^k \mathfrak{b}_n^k u_n + g^k\}),$$

$$\mathcal{K}_n^{\rho\alpha} = 2(D^\alpha \mathcal{N}_n^\rho u_n + D^\alpha h_n^\rho, D^\alpha u_n).$$

Due to Assumptions 5.1 and 5.2 we get

$$-2(D^\alpha \mathfrak{a}_n^{ij} D_i u_n, D^\alpha D_j u_n) + (D^\alpha \mathfrak{b}_n^{ik} D_i u_n, D^\alpha \mathfrak{b}_n^{jk} D_j u_n)$$

$$\leq -\lambda \sum_{i=1}^{d} |D^\alpha D_i u_n|^2 + C|u_n|_l^2$$

with a constant $C = C(K, d, d_1, l)$. Hence by standard estimates

$$\mathcal{I}_n^\alpha \leq -\frac{\lambda}{2} \sum_{i=1}^{d} |D^\alpha D_i u_n|^2 + C(|u_n|_l^2 + |f_n|_{l-1}^2 + |g_n|_l^2),$$

and by Lemma 3.4 *(i)*,

$$|\mathcal{K}_n^{\rho\alpha}| \leq C(|u_n|_l^2 + |h_n^\rho|_l^2), \quad |\mathcal{J}_n^{k\alpha}|^2 \leq C(|u_n|_l^4 + |g_n^k|_l^2 |u_n|_l^2),$$

with a constant $C = C(K, \lambda, d, d_1, l)$. Consequently, almost surely

$$|D^\alpha u_n(t)|_0^2 + \frac{\lambda}{2} \int_0^t \sum_{i=1}^{d} |D^\alpha D_i u_n|^2 \, ds \leq |D^\alpha u_{n0}|^2 + C \int_0^t |u_n|_l^2 \, dV_n$$

$$+ C \int_0^t (|f_n|_{l-1}^2 + |g_n|_l^2) \, ds + C \int_0^t |h_n^\rho|_l^2 \, d\|B_n^\rho\| + m_n^\alpha(t) \tag{5.7}$$

for every α, such that $1 \leq |\alpha| \leq l$. By virtue of (5.5) this inequality holds also for $|\alpha| = 0$. Thus summing up inequality (5.7) over all multi-indices α with $|\alpha| \leq l$ we get almost surely

$$y_n(t) := |u_n(t)|_l^2 + \frac{\lambda}{2} \int_0^t |u_n|_{l+1}^2 \, ds \leq \int_0^t |u_n|_l^2 \, dQ_n + M_n(t) + q_n(t)$$

$$\leq \int_0^t y_n \, dQ_n + M_n(t) + q_n(t) \tag{5.8}$$

for all $t \in [0, T]$ and $n \geq 1$, with

$$M_n(t) = \sum_{|\alpha| \leq l} m_n^\alpha(t),$$

$$q_n(t) = |u_{n0}|_l^2 + C \int_0^t (|f_n|_{l-1}^2 + |g_n|_l^2) \, ds + C \int_0^t |h_n^\rho|_l^2 \, d\|B_n^\rho\|(s),$$

and a constant $C = C(K, \lambda, d, d_1, l)$. Clearly,

$$d\langle m_n^\alpha \rangle = \sum_{k=1}^{d_1} |\mathcal{J}_n^{k\alpha}|^2 \, dt \leq C(|u_n|_l^4 + \gamma_n |g_n|_l^2 |u_n|_l^2) \, dQ_n,$$

so

$$d\langle M_n \rangle \leq C(|u_n|_l^4 + \gamma_n |g_n|_l^2 |u_n|_l^2) \, dQ_n \leq C(y_n^2 + \gamma_n |g_n|_l^2 y_n) \, dQ_n$$

with constants $C = C(K, \lambda, d, d_1, l)$. Hence we finish the proof of the lemma by using Assumption 5.3 and applying Lemma 3.9. $\square$

In the degenerate case, i.e., when $\lambda = 0$ in Assumption 5.1, we need to replace Assumptions 5.2 and 5.3 by somewhat stronger assumptions in order to have the conclusion of the previous lemma.

Assumption 5.4. The coefficients $\mathfrak{a}_n^{ij}$ and their derivatives in x up to order $l \vee 2$, the coefficients $\mathfrak{b}_n^{ik}$, $\mathfrak{a}_n^i$, $\mathfrak{c}_n^{i\rho}$ and their derivatives in x up to order $l \vee 1$, and the coefficients $\mathfrak{a}_n$, $\mathfrak{b}_n^k$, $\mathfrak{c}_n^\rho$ and their derivatives in x up to order l are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable real functions, in magnitude bounded by K for all $i, j = 1, ..., d$, $k = 1, ..., d_1$, $\rho = 1, ..., d_2$ and $n \geq 1$.

Assumption 5.5. We have $|u_{n0}|_l = O(n^{-\gamma})$,

$$\int_0^T |f_n|_l^2 \, ds = O(n^{-2\gamma}), \quad \int_0^T |g_n|_{l+1}^2 \, ds = O(n^{-2\gamma}),$$

$$\int_0^t |h_n(t)|_l^2 \, d\|B^\rho\|(t) = O(n^{-2\gamma}), \quad \sum_{\rho=1}^{d_2} \|B_n^\rho\|(T) = o(\ln n).$$

Theorem 5.2. *Let Assumptions 5.4, 5.5 and 5.1 (with $\lambda = 0$) hold. Let u_n be an H^{l+1}-valued weakly continuous generalized solution of* (5.1)-(5.2). *Then*

$$\sup_{t\le T}|u_n(t)|_l = O(n^{-\kappa}) \quad \textit{a.s. for each } \kappa < \gamma. \tag{5.9}$$

Proof. Let α be a multi-index such that $|\alpha| \le l$. Then, as in the proof of the previous theorem, by Itô's formula we have

$$|D^\alpha u_n(t)|^2 = |D^\alpha u_{n0}|^2 + \int_0^t \mathcal{I}_n^\alpha\, ds + \int_0^t \mathcal{K}_n^{\rho\alpha}\, dB^\rho + m_n^\alpha(t), \tag{5.10}$$

with

$$m_n^\alpha(t) = \int_0^t \mathcal{J}_n^{k\alpha}\, dW^k, \quad \mathcal{J}_n^{k\alpha} = 2(D^\alpha \mathcal{M}_n^k u_n + D^\alpha g_n^k, D^\alpha u_n),$$

$$\mathcal{I}_n^\alpha = -2(D^\alpha \mathfrak{a}_n^{ij} D_i u_n, D_j D^\alpha u_n) - 2\big(D^\beta\{(\mathfrak{a}_n^i - \mathfrak{a}_{nj}^{ij}) D_i u_n + \mathfrak{a}_n u + f_n\}, D^\alpha D^\gamma u_n\big)$$
$$+(D^\alpha \mathfrak{b}_n^{ik} D_i u_n + D^\alpha \mathfrak{b}_n^k u_n + D^\alpha g_n^k, D^\alpha \mathfrak{b}_n^{jk} D_j u_n + D^\alpha \mathfrak{b}_n^k u_n + D^\alpha g_n^k),$$
$$\mathcal{K}_n^{\rho\alpha} = 2(D^\alpha \mathcal{N}_n^\rho u_n + D^\alpha h_n^\rho, D^\alpha u_n),$$

where β and γ are multi-indices such that $\alpha = \beta + \gamma$ and $|\gamma| = 1$ if $|\alpha| \ge 1$. By [15, Lemma 2.1] and [15, Remark 2.1],

$$\mathcal{I}_n^\alpha \le C(|u_n|_l^2 + |f_n|^2 + |g_n|_{l+1}^2),$$

and by Lemma 3.4 *(i)*,

$$|\mathcal{K}_n^{\rho\alpha}| \le C(|u_n|_l^2 + |h_n^\rho|_l^2)$$

with a constant $C = C(K, d, d_1, l)$. Thus from (5.10) we get

$$|D^\alpha u_n(t)|^2 \le |D^\alpha u_{n0}|^2 + \int_0^t |u_n|_l^2\, dQ_n$$

$$+ C\int_0^t (|f_n|_l^2 + |g_n|_{l+1}^2)\, ds + C\int_0^t |h_n^\rho|_l^2\, d\|B_n^\rho\| + m_n^\alpha(t) \tag{5.11}$$

for $|\alpha| \le l$, where $Q_n(s) = C\left(s + \sum_{\rho=1}^{d_2} \|B_n^\rho\|(s)\right)$. Summing up these inequalities over α, $|\alpha| \le l$, we obtain

$$y_n(t) := |u_n(t)|_l^2 \le \int_0^t |u_n|_l^2\, dQ_n + M_n(t) + q_n(t) \tag{5.12}$$

for all $t \in [0, T]$ and $n \ge 1$, where

$$M_n(t) = \sum_{|\alpha|\le l} m_n^\alpha(t),$$

$$q_n(t) = |u_{n0}|_l^2 + C\int_0^t (|f_n|_l^2 + |g_n|_{l+1}^2)\, ds + C\int_0^t |h_n^\rho|_l^2\, d\|B_n^\rho\|,$$

and $C = C(K, d, d_1, l)$ is a constant. Hence the rest of the proof is the same as that in the proof of the previous theorem. □

6. Proof of the main theorems

To prove our main results we look for processes r_n such that

$$\sup_{t\le T}|r_n(t)|_m = O(n^{-\kappa}) \quad \text{a.s. for each } \kappa < \gamma, \tag{6.1}$$

and $v_n := u - u_n - r_n$ solves a suitable Cauchy problem of the type (5.1)-(5.2), satisfying the conditions of Theorem 5.1 or Theorem 5.2, so that we could get for each $\kappa < \gamma$

$$\sup_{t\le T}|v_n|_m = O(n^{-\kappa}) \quad \text{a.s. for each } \kappa < \gamma.$$

6.1. **Proof of Theorem 2.2.** We will carry out the strategy above in several steps, formulated as lemmas below. By a well-known result, see, e.g., [16], u is an H^{m+1}-valued strongly continuous process, and

$$\sup_{t\leq T}|u|^2_{m+1}+\int_0^T|u|^2_{m+2}\,dt<\infty \quad \text{almost surely.} \tag{6.2}$$

Moreover, we can apply Theorem 4.1 with $l=m+3$ to get

$$\sup_{t\leq T}|u_n|^2_{m+3}+\int_0^T|u_n|^2_{m+4}\,dt=O(n^{\varepsilon}) \quad \text{a.s. for any } \varepsilon>0. \tag{6.3}$$

Notice that $u-u_n$ satisfies

$$\begin{aligned}d(u-u_n)=&\{L_n(u-u_n)+\bar f_n\}\,dt+\{M^k_n(u-u_n)+\bar g^k_n\}\,dW^k\\ &+\tfrac12\{M^kM^ku+M^kg^k\}\,dt+(M^k_nu_n+g^k_n)\,d(W^k-W^k_n),\end{aligned} \tag{6.4}$$

with

$$\bar f_n:=f-f_n+(L-L_n)u, \quad \bar g^k_n:=g^k-g^k_n+(M^k-M^k_n)u.$$

Notice also that due to (6.2) and Assumption 2.5 we have

$$\int_0^T(|\bar f_n|^2_{m-1}+|\bar g_n|^2_m)\,dt=O(n^{-2\gamma}). \tag{6.5}$$

Next we rewrite equation (6.4) as an equation for

$$w_n=u-u_n-z_n, \quad \text{where } z_n=(M^k_nu_n+g^k_n)(W^k-W^k_n).$$

Note that by (6.3) and by our assumptions we have for each $\kappa<\gamma$

$$\sup_{t\leq T}|z_n(t)|^2_m+\int_0^T|z_n(t)|^2_{m+1}\,dt=O(n^{-2\kappa}) \quad \text{almost surely.} \tag{6.6}$$

Set $\mathcal{L}_n:=L_n+\frac12M^k_nM^k_n$ and recall the definition of S^{kl}_n in Remark 2.1.

Lemma 6.1. *The process w_n solves*

$$\begin{aligned}dw_n=&(\mathcal{L}_nw_n+F_n)\,dt+(M^k_nw_n+G^k_n)\,dW^k\\ &-M^k_n(M^l_nu_n+g^l_n)\,dS^{kl}_n,\end{aligned} \tag{6.7}$$

where $G^k_n=\bar g^k_n+M^k_nz_n$ and

$$\begin{aligned}F_n=&\bar f_n+\tfrac12(M^kM^k-M^k_nM^k_n)u+\tfrac12(M^kg^k-M^k_ng^k_n)\\ &-(M^k_n(L_nu_n+f_n))(W^k-W^k_n)+\mathcal{L}_nz_n.\end{aligned}$$

Proof. By using Itô's formula one can easily verify that

$$\begin{aligned}dz_n=&M^k_n(L_nu_n+f_n)(W^k-W^k_n)\,dt+M^k_n(M^l_nu_n+g^l_n)(W^k-W^k_n)\,dW^l_n\\ &+(M^k_nu_n+g^k_n)\,d(W^k-W^k_n).\end{aligned} \tag{6.8}$$

Hence

$$\begin{aligned}dw_n=&\{L_n(u-u_n)+\bar f_n+\tfrac12M^kM^ku+\tfrac12M^kg^k\\ &-M^k_n(L_nu_n+f_n)(W^k-W^k_n)\}\,dt\\ &+\{M^k_n(u-u_n)+\bar g^k_n\}\,dW^k-M^k_n(M^l_nu_n+g^l_n)(W^k-W^k_n)\,dW^l_n\\ =&\{\mathcal{L}_n(u-u_n)+\bar f_n+\tfrac12(M^kM^k-M^k_nM^k_n)u+\tfrac12(M^kg^k-M^k_ng^k_n)\}\,dt\\ &-(M^k_nL_nu_n+f_n)(W^k-W^k_n))\,dt\\ &+(M^k_nw_n+G^k_n)\,dW^k-M^k_n(M^l_nu_n+g^l_n)\,dS^{kl}_n\\ =&(\mathcal{L}_nw_n+F_n)\,dt+(M^kv_n+G^k_n)\,dW^k-M^k_n(M^l_nu_n+g^l_n)\,dS^{kl}_n,\end{aligned}$$

The lemma is proved. ◻

It is easy to show that due to (6.5), (6.3), (6.2), Assumptions 2.3, 2.4, 2.5, and 2.1 (i) we have

$$\int_0^T (|F_n|^2_{m-1} + |G_n|^2_m)\,dt = O(n^{-2\kappa}) \quad \text{(a.s.) for each } \kappa < \gamma. \tag{6.9}$$

We rewrite the last term in the right-hand side of (6.7) into symmetric and antisymmetric parts as follows:

$$\begin{aligned} M^k_n(M^l_n u_n + g^l_n)\,dS^{kl}_n &= \tfrac{1}{2}(M^k_n M^l_n + M^l_n M^k_n)u_n\,dS^{kl}_n \\ +\tfrac{1}{2}(M^k_n M^l_n - M^l_n M^k_n)u_n\,dS^{kl}_n &+ \tfrac{1}{2}M^k_n g^l_n\,d(S^{kl}_n + S^{lk}_n) + \tfrac{1}{2}(M^l_n g^k_n - M^k_n g^l_n)\,dS^{lk}_n \\ &= \tfrac{1}{2}M^k_n(M^l_n u_n + g^l_n)\,d(S^{kl}_n + S^{lk}_n) + \tfrac{1}{2}\left([M^l_n, M^k_n]u_n + M^k_n g^l_n - M^l_n g^k_n\right)\,dS^{kl}_n, \end{aligned}$$

where $[A, B] = BA - AB$. Thus using Remark 2.1 we get

$$\begin{aligned} M^k_n(M^l_n u_n + g^l_n)\,dS^{kl}_n &= -\tfrac{1}{2}M^k_n(M^l_n u_n + g^l_n)\,dq^{kl}_n \\ +\tfrac{1}{2}M^k_n(M^l_n u_n + g^l_n)\,d(R^{kl}_n + R^{lk}_n) &+ \tfrac{1}{2}\left([M^l_n, M^k_n]u_n + M^k_n g^l_n - M^l_n g^k_n\right)\,dS^{kl}_n \\ &= -\tfrac{1}{2}M^k_n(M^l_n u_n + g^l_n)\,dq^{kl}_n + \tfrac{1}{2}\left(M^l_n \diamond M^k_n u_n + M^k g^l_n + M^l_n g^k_n\right)\,dR^{kl}_n \\ &\quad + \tfrac{1}{2}\left([M^l_n, M^k_n]u_n + M^k_n g^l_n - M^l_n g^k_n\right)\,dS^{kl}_n, \end{aligned} \tag{6.10}$$

where we use the notation $A \diamond B = BA + AB$ for linear operators A and B. Thus equation (6.7) can be rewritten as follows.

Lemma 6.2. *The process w_n solves*

$$\begin{aligned} dw_n = (\mathcal{L}_n w_n + F_n)\,dt &+ (\tfrac{1}{2}[M^l_n, M^k_n]w_n + H^{kl}_n)\,dS^{kl}_n + (M^k_n w_n + \bar{G}^k_n)\,dW^k \\ &+ \tfrac{1}{2}M^k_n(M^l_n u_n + g^l_n)\,dq^{kl}_n + \tfrac{1}{2}\left([M^k, M^l]u + M^l g^k - M^k g^l\right)\,dS^{kl}_n, \end{aligned} \tag{6.11}$$

where

$$\begin{aligned} H^{kl}_n =& \tfrac{1}{2}[M^l_n, M^k_n](M^r_n u_n + g^r_n)(W^r - W^r_n) \\ &+ \tfrac{1}{2}([M^k_n, M^l_n] - [M^k, M^l])u + \tfrac{1}{2}(M^l_n g^k_n - M^l g^k + M^k g^l - M^k_n g^l_n), \\ \bar{G}^k_n =& G^k_n - \tfrac{1}{2}\left(M^k_n \diamond M^l_n u_n + M^l_n g^k_n + M^k_n g^l_n\right)(W^l - W^l_n) \end{aligned}$$

Proof. Plugging (6.10) into (6.7) we get

$$\begin{aligned} dw_n =& (\mathcal{L}_n w_n + F_n)\,dt + (M^k_n w_n + \bar{G}^k_n)\,dW^k \\ &+ \tfrac{1}{2}M^k_n(M^l_n u_n + g^l_n)\,dq^{kl}_n - \tfrac{1}{2}\{[M^l_n, M^k_n]u_n + M^k_n g^l_n - M^l_n g^k_n\}\,dS^{kl}_n \\ =& (\mathcal{L}_n w_n + F_n)\,dt + (M^k_n w_n + \bar{G}^k_n)\,dW^k \\ &+ \tfrac{1}{2}M^k_n(M^l_n u_n + g^l_n)\,dq^{kl}_n + \left(\tfrac{1}{2}[M^l_n, M^k_n]w_n + M^l_n g^k_n - M^k_n g^l_n\right)\,dS^{kl}_n \\ &- \tfrac{1}{2}[M^l_n, M^k_n]u\,dS^{kl}_n + \tfrac{1}{2}[M^l_n, M^k_n](M^r_n u_n + g^r_n)(W^r - W^r_n)\,dS^{kl}_n \\ =& (\mathcal{L}_n w_n + F_n)\,dt + (\tfrac{1}{2}[M^l_n, M^k_n]w_n + H^{kl}_n)\,dS^{kl}_n + (M^k_n w_n + \bar{G}^k_n)\,dW^k \\ &+ \tfrac{1}{2}M^k_n(M^l_n u_n + g^l_n)\,dq^{kl}_n + \tfrac{1}{2}\{[M^k, M^l]u + M^l g^k - M^k g^l\}\,dS^{kl}_n. \end{aligned}$$

□

In the same way as (6.9) is proved, we can easily get

$$\int_0^T |\bar{G}_n|^2_{m-1}\,dt = O(n^{-2\kappa}), \quad \sup_{t\le T}|H^{kl}_n(t)|^2_m = O(n^{-2\kappa}) \quad \text{for } \kappa < \gamma \tag{6.12}$$

almost surely for all $k, l = 1, ..., d_1$. Finally we rewrite (6.11) as an equation for $v_n = w_n - r_n$, where

$$r_n = \tfrac{1}{2}M^k_n(M^l_n u_n + g^l_n)q^{kl}_n + \tfrac{1}{2}\{[M^k, M^l]u + M^k g^l - M^l g^k\}S^{kl}_n.$$

Notice that by (6.3) and Remark 2.1, r_n satisfies (6.6) in place of z_n.

Lemma 6.3. *The process v_n solves*

$$
\begin{aligned}
dv_n =&(\mathcal{L}_n v_n + \widetilde{F}_n)\,dt + (\tfrac{1}{2}[M_n^k, M_n^l]v_n + \widetilde{H}_n^{kl})\,dS_n^{kl} + (M_n^k v_n + \widetilde{G}_n^k)\,dW^k \\
&- \tfrac{1}{2}M_n^k M_n^l (M_n^r u_n + g_n^r)(W^k - W_n^k)\,dB_n^{lr},
\end{aligned}
\tag{6.13}
$$

where B_n^{lr} is as in (2.3) *and*

$$
\begin{aligned}
\tilde{F}_n =&F_n + \mathcal{L}_n r_n - \tfrac{1}{2}M_n^k M_n^l (L_n u_n + f_n) q_n^{kl} - \tfrac{1}{2}[M^k, M^l](\mathcal{L}u + \tfrac{1}{2}M^r g^r + f)S_n^{kl}, \\
\widetilde{G}_n^k =&\bar{G}_n + M_n^k r_n - \tfrac{1}{2}[M^r, M^l](M^k u + g^k)S_n^{rl}, \\
\widetilde{H}_n^{kl} =&H_n^{kl} + \tfrac{1}{2}[M_n^l, M_n^k]r_n.
\end{aligned}
$$

Proof. Indeed,

$$
\begin{aligned}
dv_n =&(\mathcal{L}_n v_n + F_n)\,dt + (\tfrac{1}{2}[M_n^l, M_n^k]v_n + \tilde{H}_n^{kl})\,dS_n^{kl} + (M_n^k v_n + \bar{G}_n^k)\,dW^k \\
&+ \mathcal{L}_n r_n\,dt + M_n^k r_n\,dW^k \\
&- \tfrac{1}{2}M_n^k M_n^l (L_n u_n + f_n) q_n^{kl}\,dt - \tfrac{1}{2}M_n^k M_n^l (M_n^r u_n + g_n^r) q_n^{kl}\,dW_n^r \\
&- \tfrac{1}{2}[M_n^k, M_n^l](\mathcal{L}u + \tfrac{1}{2}M^r g^r + f)S_n^{kl}\,dt - \tfrac{1}{2}[M_n^k, M_n^l](M^r u + g^r)S_n^{kl}\,dW^r \\
&= (\mathcal{L}v_n + \tilde{F}_n)\,dt + (\tfrac{1}{2}[M_n^l, M_n^k]v_n + \tilde{H}_n^{kl})\,dS_n^{kl} + (M_n^k v_n + \tilde{G}_n^k)\,dW^k \\
&- \tfrac{1}{2}M_n^k M_n^l (M_n^r u_n + g_n^r)(W^k - W_n^k)\,dB_n^{lr}.
\end{aligned}
$$

□

Making use of (6.9), (6.12) and (2.5), we easily obtain that for $\kappa < \gamma$

$$
\int_0^T (|\tilde{F}_n|_{m-1}^2 + |\tilde{G}_n^k|_m^2)\,dt = O(n^{-2\kappa}), \quad \sup_{t\le T} |\tilde{H}_n^{kl}(t)|_m^2 = O(n^{-2\gamma})
\tag{6.14}
$$

almost surely for $k, l = 1, ..., d_1$. Hence we finish the proof of the theorem by applying Theorem 5.1 with $l = m$ to equation (6.13) and using (6.1) for z_n and r_n. □

6.2. **Proof of Theorem 2.4.** We follow the proof of Theorem 2.2 with the necessary changes. By a well-known theorem on degenerate stochastic PDEs from [15], u is an H^{m+2}-valued weakly continuous process, and by Theorem 4.3 with $l = m + 4$ we have

$$
\sup_{t\le T} |u_n|_{m+4}^2 = O(n^{\varepsilon}) \quad \text{a.s. for } \varepsilon > 0.
\tag{6.15}
$$

Clearly, $u - u_n$ satisfies equation (6.4), and

$$
\int_0^T (|\bar{f}_n|_m^2 + |\bar{g}_n|_{m+1}^2)\,dt = O(n^{-2\gamma}).
\tag{6.16}
$$

Moreover, Lemmas 6.1, 6.2 and 6.3 remain valid, and due to (6.15), (6.16) and our assumptions, we have for each $\kappa < \gamma$

$$
\int_0^T (|F_n|_m^2 + |G_n|_{m+1}^2)\,dt = O(n^{-2\kappa}),
\tag{6.17}
$$

$$
\int_0^T |\bar{G}_n|_{m+1}^2\,dt = O(n^{-2\kappa}), \quad \sup_{t\le T} |H_n^{kl}(t)|_m^2 = O(n^{-2\kappa}),
\tag{6.18}
$$

$$
\int_0^T |\tilde{F}_n|_m^2 + |\tilde{G}_n|_{m+1}^2\,dt = O(n^{-2\kappa}), \quad \sup_{t\le T} |\tilde{H}_n^{kl}(t)|_m^2 = O(n^{-2\kappa})
\tag{6.19}
$$

almost surely for $k, l = 1, ..., d_1$. Note also that r_n and z_n satisfy (6.1). Hence we finish the proof of the theorem by applying Theorem 5.2 with $l = m$ to equation (6.13). □

Now we prove our main results in the case when the coefficients and the free terms depend on t.

6.3. **Proof of Theorem 2.5.** We follow the proof of Theorem 2.2 with the necessary changes. As before, (6.2) and (6.3) hold. Now $u - u_n$ satisfies equation (6.4) with an additional term,

$$\frac{1}{2}\sum_{k=1}^{d_1}\big(M^{k(k)}u + g^{k(k)}\big)\,dt,$$

added to the right-hand side of (6.4). Thus to get the analogue of Lemma 6.1 we set

$$N_n = \frac{1}{2}\sum_{k=1}^{d_1} M_n^{k(k)}, \quad \bar{\mathcal{L}}_n = \mathcal{L}_n + N_n = L_n + \frac{1}{2}\Big(M_n^k M_n^k + \sum_{k=1}^{d_1} M_n^{k(k)}\Big),$$

$$M_n^{k(l)} = b_n^{ik(l)} D_i + b_n^{k(l)}, \quad M^{kl} = b^{ik(l)} D_i + b^{k(l)}$$

for $k = 1, ..., d_1$, $l = 0, ..., d_1$ and $n \geq 1$. Then for

$$w_n = u - u_n - z_n, \quad z_n = (M_n^k u_n + g_n^k)(W^k - W_n^k) \tag{6.20}$$

the corresponding lemma reads as follows.

Lemma 6.4. *The process w_n solves*

$$\begin{aligned} dw_n =& (\bar{\mathcal{L}}_n w_n + \bar{F}_n)\,dt + (M_n^k w_n + G_n^k)\,dW^k \\ & - M_n^k(M_n^l u_n + g_n^l)\,dS_n^{kl} - (M_n^{k(l)} u_n + g_n^{k(l)})\,dS_n^{kl}, \end{aligned}$$

where

$$\begin{aligned} \bar{F}_n =& F_n + \frac{1}{2}\sum_{k=1}^{d_1}(M^{k(k)} - M_n^{k(k)})u + \frac{1}{2}\sum_{k=1}^{d_1}(g^{k(k)} - g_n^{k(k)}) \\ & - (M^{k(0)}u_n + g^{k(0)})(W^k - W_n^k) + N_n z_n, \end{aligned} \tag{6.21}$$

and F_n and G_n are defined in Lemma 6.1.

Proof. We need only notice that for z_n equation (6.8) holds with a new term,

$$(M_n^{k(0)}u_n + g^{k(0)})(W^k - W_n^k)\,dt + (M_n^{k(l)}u_n + g_n^{k(i)})(W^k - W_n^k)\,dW_n^l?,$$

added to its right-hand side. □

Hence we get the following modification of Lemma 6.2

Lemma 6.5. *The process w_n solves*

$$\begin{aligned} dw_n =& (\bar{\mathcal{L}}_n w_n + \bar{F}_n)\,dt + (\tfrac{1}{2}([M_n^l, M_n^k] + M_n^{k(l)})w_n + \bar{H}_n^{kl})\,dS_n^{kl} \\ & + (M_n^k w_n + \bar{G}_n^k)\,dW^k + \tfrac{1}{2}M_n^k(M_n^l u_n + g_n^l)\,dq_n^{kl} \\ & + \tfrac{1}{2}\Big(([M^k, M^l] - M^{k(l)})u + M^l g^k - M^k g^l - g^{k(l)}\Big)\,dS_n^{kl}, \end{aligned} \tag{6.22}$$

where

$$\begin{aligned} \bar{H}_n^{kl} =& H_n^{kl} + \tfrac{1}{2}M_n^{k(l)}(M_n^r u_n + g_n^r)(W^r - W_n^r) \\ & + \tfrac{1}{2}(M^{k(l)} - M_n^{k(l)})u + \tfrac{1}{2}(g^{k(l)} - g_n^{k(l)}), \end{aligned} \tag{6.23}$$

and H_n^{kl} and $\bar{G}_n^k$ are defined in Lemma 6.2.

Now we rewrite equation (6.22) as an equation for $\bar{v}_n := w_n - \bar{r}_n$, where

$$\bar{r}_n = r_n - \tfrac{1}{2}(M^{k(l)}u + g^{k(l)})S_n^{kl} \tag{6.24}$$

$$= \tfrac{1}{2}M_n^k(M_n^l u_n + g_n^l)q_n^{kl} + \tfrac{1}{2}(([M^k, M^l] - M^{k(l)})u + M^l g^k - M^k g^l - g^{k(l)})S_n^{kl}.$$

To this end we set

$$\tilde{f} = f + \tfrac{1}{2}M^k g^k + \tfrac{1}{2}\sum_{k=1}^{d_1} g^{k(k)},$$

and notice that

$$\begin{aligned}dM^k_n(M^l_n u_n + g^l_n) =&M^k_n M^l_n(L_n u_n + f_n)\,dt + M^k_n M^l_n(M^j_n u_n + g^j_n)\,dW^j_n\\&+ T^{kl0}_n\,dt + T^{klj}_n\,dW^j_n,\end{aligned}$$

where

$$\begin{aligned}T^{kl0}_n &= (M^{k(0)}_n M^l_n + M^k_n M^{l(0)}_n)u_n + M^{k(0)}_n g^l_n + M^k_n g^{l(0)}_n,\\T^{klj}_n &= (M^{k(j)}_n M^l_n + M^k_n M^{l(j)}_n)u_n + M^{k(j)}_n g^l_n + M^k_n g^{l(j)}_n.\end{aligned}$$

Similarly,

$$\begin{aligned}d[M^k, M^l]u =&[M^k, M^l](\mathcal{L}u + f + \tfrac{1}{2}M^j g^j)\,dt + [M^k, M^l](M^j + g^j)\,dW^j\\&+ P^{kl0}\,dt + P^{klj}\,dW^j,\end{aligned}$$

$$d(-M^{k(l)}u + M^l g^k - M^k g^l - g^{k(l)}) = U^{kl0}\,dt + U^{klj}\,dW^j,$$

where $\mathcal{L} = L + \frac{1}{2}M^k M^k$,

$$\begin{aligned}P^{kl0} =&[M^k, M^{l(0)}]u + [M^{k(0)}, M^l]u + \sum_{j=1}^{d_1}[M^{k(j)}, M^{l(j)}]u\\&+ ([M^k, M^{l(j)}] + [M^{k(j)}, M^l])(M^j u + g^j)\\&+ \tfrac{1}{2}\sum_{j=1}^{d_1}[M^k, M^l](M^{j(j)}u + g^{j(j)}),\\P^{klj} =&([M^k, M^{l(j)}] + [M^{k(j)}, M^l])u,\\U^{kl0} =&- M^{k(l0)}u - M^{k(l)}(\bar{\mathcal{L}}u + \tilde{f}) - M^{k(lj)}(M^j u + g^j)\\&+ M^{l(0)}g^k + M^l g^{k(0)} + M^{l(j)}g^{k(j)}\\&- M^{k(0)}g^l - M^k g^{l(0)} - M^{k(j)}g^{l(j)} - g^{k(l0)},\\U^{klj} =&- M^{k(lj)}u - M^{k(l)}(M^j u + g^j) + M^{l(j)}g^k + M^l g^{k(j)}\\&- M^{k(j)}g^l - M^k g^{l(j)} - g^{k(lj)}.\end{aligned}$$

Let $\tilde{F}_n$, $\tilde{G}^k_n$ and $\tilde{H}^{kl}_n$ be defined now as in Lemma 6.3, but with F_n and G^k_n replaced there by $\bar{F}_n$ and $\bar{G}^k_n$ in (6.21) and (6.23), respectively.

Thus we have the following modification of Lemma 6.3.

Lemma 6.6. *The process $\bar{v}_n = w_n - \bar{r}_n$ solves*

$$\begin{aligned}d\bar{v}_n =&(\bar{\mathcal{L}}_n\bar{v}_n + \hat{F}_n)\,dt + (\tfrac{1}{2}([M^k_n, M^l_n] + M^{k(l)})\bar{v}_n + \hat{H}^{kl}_n)\,dS^{kl}_n\\&+ (M^k_n v_n + \hat{G}^k_n)\,dW^k - \tfrac{1}{2}T^{klj}(W^k - W^k_n)\,dB^{lr}_n,\end{aligned}\tag{6.25}$$

where

$$\begin{aligned}\hat{F}_n =&\tilde{F}_n + N_n r_n - \tfrac{1}{2}\bar{\mathcal{L}}_n(M^{k(l)}u + g^{k(l)})S^{kl}_n\\&- \tfrac{1}{2}T^{kl0}q^{kl}_n - \tfrac{1}{2}(P^{kl0}_n + U^{kl0}_n)S^{kl}_n,\\\hat{G}^k_n =&\widetilde{G}_n - \tfrac{1}{2}M^k_n(M^{j(l)}u + g^{j(l)})S^{jl}_n - \tfrac{1}{2}(P^{jlk} + U^{jlk})S^{jl}_n\\\hat{H}^{kl}_n =&\tilde{H}^{kl}_n + \tfrac{1}{2}M^{k(l)}r_n - \tfrac{1}{4}([M^l_n, M^k_n] + M^{kl}_n)(M^{k(l)}u + g^{k(l)})S^{kl}_n.\end{aligned}$$

We can verify that (6.14) holds with $\hat{F}_n$, $\hat{G}^k_n$ and $\hat{H}^{kl}$ in place of $\tilde{F}_n$, $\tilde{G}^k_n$ and $\tilde{H}^{kl}$, respectively. We can also see that z_n and $\bar{r}_n$ satisfy (6.6). Hence we finish the proof by applying Theorem 5.1 with $l = m$ to equation (6.25).

6.4. **Proof of Theorem 2.6.** We get Lemma 6.6 in the same way as Lemma 6.3 is proved, and we can also see that

$$\int_0^T (|\hat{F}_n|_m^2 + |\hat{G}_n^k|_{m+1}^2)\, dt = O(n^{-2\kappa}), \quad \sup_{t\le T} |\hat{H}_n^{kl}(t)|_m^2 = O(n^{-2\kappa})$$

for each $\kappa < \gamma$, almost surely for $k, l = 1, ..., d_1$, where $\hat{F}_n$, $\hat{G}_n^k$ and $\hat{H}_n^{kl}$ are defined in Lemma 6.3. We can also verify that for z_n and $\bar{r}_n$, defined in (6.20) and (6.24), we have

$$\sup_{t\le T} |z_n(t)|_m + \sup_{t\le T} |\bar{r}_n(t)|_m = O(n^{-\kappa}) \quad \text{for each } \kappa < \gamma.$$

Hence we obtain the theorem by applying Theorem 5.2 with $l = m$ to equation (6.25).

Acknowledgments. The first author is grateful to the organisers of the "Seventh Seminar on Stochastic Analysis, Random Fields and Applications" for the invitation and for the possibility of giving a talk on the results of this paper, which were obtained while the authors were staying in the Isaac Newton Institute of Cambridge University in April 2010. The second author is also grateful to the University of Edinburgh's School of Mathematics for their kind hospitality during several visits.

References

[1] P. Acquistapace and B. Terreni, *An approach to Ito linear equations in Hilbert spaces by approximation of white noise with coloured noise*, Stochastic Anal. Appl. **2** (1984), no. 2, 131–186.
[2] V. Bally, A. Millet and M. Sanz-Solé, *Approximation and support theorem in Hölder norm for parabolic stochastic partial differential equations*, Ann. Prob. **23** (1995), no. 1, 178–222.
[3] Z. Brzeźniak, M. Capinski and F. Flandoli, *A convergence result for stochastic partial differential equations*, Stochastics **24** (1988), no. 4, 423–445.
[4] Z. Brzeźniak and F. Flandoli, *Almost sure approximation of Wong-Zakai type for stochastic partial differential equations*, Stochastic Process. Appl. **55** (1995), no. 2, 329–358
[5] P. Friz and H. Oberhauser, *Rough path limits of the Wong-Zakai type with a modified drift term*, J. Funct. Anal. **256** (2009), no. 10, 3236–3256.
[6] P. Friz and H. Oberhauser, *Rough path stability of SPDEs arising in non-linear filtering*, preprint arXiv:1005.1781 (2010), 17 pp.
[7] P. Friz and H. Oberhauser, *Rough path stability of (semi-) linear SPDEs*, preprint (2011).
[8] I. Gyöngy, *On the approximation of stochastic partial differential equations I*, Stochastics **25** (1988), no. 2, 59–85.
[9] I. Gyöngy, *On the approximation of stochastic partial differential equations II*, Stochastics **26** (1989), no. 3, 129–164.
[10] I. Gyöngy, *Introduction to Stochastic Partial Differential Equations*, in preparation for publication.
[11] I. Gyöngy and N. V. Krylov, *On the Rate of Convergence of Splitting-up Approximations for SPDEs*, in: Stochastic Inequalities and Applications, Progr. Prob. **56**, 301–321, Birkhäuser, Basel, 2003.
[12] I. Gyöngy and A. Shmatkov, *Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations*, Appl. Math. Optim. **54** (2006), no. 3, 315–341.
[13] I. Gyöngy and A. Shmatkov, *On the rate of convergence of Wong-Zakai approximations*, Preprint.
[14] E. Hausenblas, *Wong-Zakai type approximation of SPDEs of Lévy noise*, Acta Appl. Math. **98** (2007), no. 2, 99–134.
[15] N.V. Krylov and B.L. Rozovskiĭ, *Characteristics of second-order degenerate parabolic Itô equations*, Trudy sem. Petrovsk. **8** (1982), 153–168 (in Russian); English translation in *J. Soviet Math.* **32**, (1986), 336–348.
[16] B.L. Rozovskii, *Stochastic Evolution Systems. Linear Theory and Applications to Nonlinear Filtering, Mathematics and its Applications (Soviet Series)* ***35****, Kluwer, Dordrecht, 1990.*
[17] A. Shmatkov, *The Rate of Convergence of Wong-Zakai Approximations for SDEs and SPDEs*, PhD dissertation, University of Edinburgh, 2006.
[18] G. Tessitore and J. Zabczyk, *Wong-Zakai approximations of stochastic evolution equations*, J. Evol. Equ. **6** (2006), no. 4, 621–655.

School of Mathematics and Maxwell Institute, University of Edinburgh, King's Buildings, EH9 3JZ Edinburgh, United Kingdom
E-mail address: gyongy@maths.ed.ac.uk

Department of Mathematics, The University of Texas at Austin, 1 University Station, C1200, Austin, TX 78712-1202, United States of America
E-mail address: stinga@math.utexas.edu